

上海市研究生教育专项经费资助项目

凸分析与非光滑分析

胡毓达 孟志青 著

上海科学技术出版社

图书在版编目(CIP)数据

凸分析与非光滑分析/胡毓达,孟志青著. —上海:
上海科学技术出版社, 2000. 9

上海研究生教育用书

ISBN 7-5323-5525-X

I. 凸... II. ①胡...②孟... III. 凸分析-研究生-
教材 IV. 0174.13

中国版本图书馆 CIP 数据核字(2000)第 41258 号

责任编辑 赵序明

上海科学技术出版社出版发行

(上海瑞金二路 450 号 邮政编码 200020)

上海新华印刷厂印刷 新华书店上海发行所经销

2000 年 9 月第 1 版 2000 年 9 月第 1 次印刷

开本 850×1156 1/32 印张 12 插页 4 字数 312 000

印数 1-1 500 定价: 25.00 元

本书如有缺页、错装或坏损等严重质量问题,
请向本社出版科联系调换

内 容 提 要

本书论述凸分析与非光滑分析的基本理论,特别是较系统地介绍了有关集值映射的广义微分的研究进展.全书共8章:第1章至第3章是凸分析部分,包括凸集、凸函数和凸映射以及凸函数的次微分;第4章概述非光滑的Lipschitz函数的广义次微分;第5章则利用切锥引进并讨论函数的切导数和切微分;第6章至第8章研究集值映射的广义微分,其中第6章考虑锥凸集值映射的锥次微分;第7章进而讨论一般集值映射的锥弱次微分;第8章则从另一角度研究集值映射的切导数和切微分.

本书适用于作大学硕士和博士研究生的教材或主要教学参考书,也可供大学本科高年级学生选读.同时,它也是从事现代分析、数学规划、向量极值、最优控制和数理经济等领域研究的学者的重要参考书.

前 言

凸分析和非光滑分析是 20 世纪 60 年代至 80 年代相继发展形成的现代数学分支. 作为描述和解决非光滑问题的有力工具, 它们在非线形最优化、多目标决策、最优控制、对策论、变分学、逼近理论以及数理经济等领域有着广泛的应用. 事实上, 凸分析主要研究的凸函数(凸泛函)在通常意义下是非光滑的, 因此从这一意义上说, 凸分析也是非光滑分析的组成部分和重要基础, 而非光滑分析则是凸分析研究的延伸和发展.

本书是我自 1986 年至 1998 年间在上海交通大学先后为应用数学专业的硕士和博士研究生讲授凸分析、非光滑分析和现代分析课的基础上, 经过选择整理写成的. 全书脱稿后, 于 2000 年 2 月至 6 月在温州大学对该校和温州师范学院的数学教师以及访问学者试讲了部分内容, 并稍作修改.

书中的第 1 章至第 3 章是相对成熟的凸分析部分, 第 2 章中有关映射的拟凸性的内容, 主要是我和学生胡一凡博士以及访问学者凌晨副教授合作的工作. 第 4 章关于 Lipschitz 函数的广义次微分和第 5 章函数的切微分, 部分取材于 F. H. Clarke 的著作^[11]. 第 5 章和第 8 章的部分内容, 参考了 J. P. Aubin 和 I. Ekeland 的书^[16], 以及 B. S. Mordukhovich 的论文^[28]. 第 6 章至第 8 章是集值映射的广义微分部分, 其基本概念先是由我和学生徐永明博士开始引进, 而后则是和孟志青博士共同进行了较系统研究的结果. 本书的第 1 章至第 3 章由我执笔撰写. 1994 年至 1995 年间, 湘潭大学数学系孟志青副教授作为访问学者来我处进修, 他除了与我合作进行有关集值映射的广义微分的研究之外, 还曾在我的非光滑分析课中分讲了部分内容. 为此, 本书的第 4 章至第 8 章先

由他整理出初稿,经我们多次讨论调整,最后由我作修改、补充和统一定稿。

阅读本书一般需要有泛函分析的基本知识,特别是从第 5 章以后,读者最好能熟悉非线性泛函的某些近代理论,同时对其中一些比较简洁的推证要进行仔细地研读。

在本书的写作过程中,陆晋奎教授、王晓敏副教授和杨雷博士曾协助我补充和查对了有关资料,丁鸿生讲师和于丽英博士则为打印本书的手稿付出了辛勤的劳动。在此,一并致以衷心感谢!

限于作者水平,书中难免存在不妥和谬误之处,恳望专家和读者不吝批评指正。

胡毓达

2000 年 1 月 1 日于上海交通大学

2000 年 7 月 7 日补写于温州大学

符号说明

V, W	线性空间
L	线性子空间
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$	线性拓扑空间
$\mathcal{X}^*, \mathcal{Y}^*, \mathcal{Z}^*$	线性拓扑空间 $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ 的对偶空间
\mathcal{H}	Hilbert 空间
$\mathcal{B}, \mathcal{D}, \mathcal{B}_1, \mathcal{B}_2$	Banach 空间
\mathcal{B}^*	Banach 空间 \mathcal{B} 的对偶空间
R^n, R^m	n 维, m 维 Euclid 空间
R	实数域
R_+	非负实数域
S, T, M	集合
$\text{co}S$	集合 S 的凸包
$\text{aff}S$	集合 S 的仿射包
$\text{lin}S$	集合 S 的线性包
S^i	集合 S 的代数内部
S^r	集合 S 的相对代数内部
S^a	集合 S 的代数闭包
S^b	集合 S 的代数边界
$\text{int}S$	集合 S 的内部
$\text{ri}S$	集合 S 的相对内部
$\text{cl}S, \bar{S}$	集合 S 的闭包
∂S	集合 S 的边界
$d \perp S$	Banach 空间中向量 d 垂直于集合 S
S^\perp	Banach 空间中集合 S 的垂直集

H	超平面
H_+^c, H_-^c	由超平面 H 界定的闭半空间
H_+^o, H_-^o	由超平面 H 界定的开半空间
$U(x)$	线性拓扑空间中点 x 的邻域
$U, U(O)$	线性拓扑空间中原点的邻域
$U(x, \delta)$	Banach 空间中点 x 的 δ -邻域
$B_\delta(x)$	Banach 空间中以点 x 为中心, δ 为半径的 超球
B	Banach 空间中的开单位球(或闭单位球)
(x^1, x^2)	线性空间中点 x^1 与 x^2 之间的开线段
$[x^1, x^2]$	线性空间中点 x^1 与 x^2 之间的闭线段
$(x^1, x^2], [x^1, x^2)$	线性空间中点 x^1 与 x^2 之间的半开半闭线段
$\ \cdot\ _\infty, \ \cdot\ $	Banach 空间上的范数
$\ \cdot\ _p$	Euclid 空间上的 p -范数
$\ M\ _{m \times n}$	$m \times n$ 阶矩阵 M 的范数
K	线性空间中的锥
K^*	锥 K 的对偶锥
K^{**}	锥 K 的双对偶锥
K^+	锥 K 的严格对偶锥
K^-	锥 K 的负对偶锥
R_+^n	n 维 Euclid 空间中的非负锥
S^-	集合 S 的负对偶集(极化集)
$\text{cone} S$	集合 S 的锥包
f, g	线性空间上的实值函数(实泛函)
f	线性空间上的实值向量函数
$\langle x^*, x \rangle$	线性空间上线性泛函 x^* 在点 x 处的值
φ	线性拓扑空间上的映射
φ^*	线性映射 φ 的对偶算子
ψ	线性拓扑(Banach)空间上的集值映射

L	线性空间到线性空间的线性映射
$L(\mathcal{X}, \mathcal{Y})$	线性拓扑空间 \mathcal{X} 到 \mathcal{Y} 的连续线性映射(算子)集合
$\text{dom} f$	广义实值函数 f 的有效域
$\text{epi} f$	实值函数 f 的上图象
$\text{hyp} f$	实值函数 f 的下图象
$\text{co}(f)$	实值函数 f 的凸包(函数)
$\text{cl}(f)$	实值函数 f 的闭包(函数)
$\liminf_{x \rightarrow x^0} f(x)$	实值函数 f 在点 x^0 处的下极限
$\limsup_{x \rightarrow x^0} f(x)$	实值函数 f 在点 x^0 处的上极限
$\text{dom} \psi$	集值映射 ψ 的有效域
$\text{graph} \psi$	集值映射 ψ 的图象
$K\text{-epi} \psi$	集值映射 ψ 的 K -上图象
$\text{Ker} \psi$	集值映射 ψ 的核
$\text{Im} \psi$	集值映射 ψ 的象
$H_S(f, c)$	集合 S 上函数 f 关于 c 的水平集
$H_S(\varphi, c)_K$	集合 S 上映射 φ 关于 c 的 K -水平集
$K\text{-bou}\{y^1, \dots, y^m\}$	线性拓扑空间中向量组 $\{y^1, \dots, y^m\}$ 的 K -有界集
f^*	凸函数 f 的共轭函数
f^{**}	凸函数 f 的二次共轭函数
g^*	凹函数 g 的共轭函数
$d_S, d_S(x)$	Banach 空间中集合 S 上的距离函数
$\mu_S, \mu_S(x)$	线性拓扑空间中凸集 S 上的 Minkowski 函数(泛函)
$\mu_a, \mu_a(y, \omega)$	$\mathcal{Y} \times \text{int} K$ 上关于 $a \in \mathcal{Y}$ 的 Minkowski 泛函
$\delta_S, \delta_S(x)$	线性空间中凸集 S 上的指示函数
$\sigma_S, \sigma_S(x)$	线性空间的对偶空间中集合 S 的支撑函数
$l, l(x)$	线性空间上的仿射函数

$T_S(\mathbf{x}^0)$	线性拓扑空间中凸集 S 在点 \mathbf{x}^0 处的切锥
$N_S(\mathbf{x}^0)$	线性拓扑空间中凸集 S 在点 \mathbf{x}^0 处的法锥
$T_S^0(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的 G -切锥 (广义切锥)
$N_S^0(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的 G -法锥 (广义法锥)
$Q_S(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的相依锥
$T_S^H(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的 H -切锥 (超切锥)
$T_S^D(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的 D -切锥 (相依切锥)
$T_S^C(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的 C -切锥 (Clarke 切锥)
$N_S^C(\mathbf{x}^0)$	Banach 空间中集合 S 在点 \mathbf{x}^0 处的 C -法锥 (Clarke 法锥)
$G_{K\text{-epi}\psi}(\mathbf{x}^0, \mathbf{y}^0)$	集合 $K\text{-epi } \psi$ 在点 $(\mathbf{x}^0, \mathbf{y}^0)$ 处的 G -切锥
$N_S^M(\mathbf{x}^0)$	集合 S 在点 \mathbf{x}^0 处的 M -法锥
$N_S^f(\mathbf{x}^0, \epsilon)$	集合 S 在点 \mathbf{x}^0 处的 F -法锥 (Frechet- ϵ 法锥)
$N_{K\text{-epi}\psi}(\mathbf{x}^0, \mathbf{y}^0)$	集值映射 ψ 在点 $(\mathbf{x}^0, \mathbf{y}^0)$ 处的法锥
$\mathcal{E}(Y, K)$	线性拓扑空间中集合 Y 的 K -有效点集
$\mathcal{E}_w(Y, K)$	线性拓扑空间中集合 Y 的 K -弱有效点集
$\nabla f(\mathbf{x}^0)$	Euclid 空间上实值函数 f 在点 \mathbf{x}^0 处的梯度
$\nabla^2 f(\mathbf{x}^0)$	Euclid 空间上实值函数 f 在点 \mathbf{x}^0 处的 Hesse 矩阵
$Jf(\mathbf{x}^0)$	Euclid 空间上实值向量函数 f 在点 \mathbf{x}^0 处的 Jacobi 矩阵
$f'(\mathbf{x}^0)$	Banach 空间上实值函数 f 在点 \mathbf{x}^0 处的 Fréchet 导数

φ'_{x^0}	线性拓扑空间上映射 φ 在点 x^0 处的 Gâteaux 导数
$D_S\varphi(x^0)$	Banach 空间上映射 φ 在点 x^0 处的严格导数
$f'_+(x^0; d)$	实值函数 f 在点 x^0 处沿方向 d 的右方向导数
$f'_-(x^0; d)$	实值函数 f 在点 x^0 处沿方向 d 的左方向导数
$f^d(x^0; d)$	实值函数 f 在点 x^0 处沿方向 d 的方向导数
$f^G(x^0; d)$	Lipschitz 函数 f 在点 x^0 处沿方向 d 的 G -方向导数(广义方向导数)
$f^D(x^0; d)$	正常函数 f 在点 x^0 处沿方向 d 的 D -切导数
$f^C(x^0; d)$	正常函数 f 在点 x^0 处沿方向 d 的 C -切导数(上图象导数)
$f^H(x^0; d)$	实值函数 f 在点 x^0 处沿方向 d 的 H -切导数
$\partial f(x^0)$	凸函数 f 在点 x^0 处的次微分
$\partial^0 f(x^0)$	函数 f 在点 x^0 处的 G -次微分
$\partial_1^0 f(x_1, x_2)$	Lipschitz 函数 f 在点 x_1 处的 G -偏次微分
$\partial_2^0 f(x_1, x_2)$	Lipschitz 函数 f 在点 x_2 处的 G -偏次微分
$\partial^0 f(x^0)$	Lipschitz 向量函数 f 在点 x^0 处的 G -次微分
$\partial^C f(x^0)$	正常函数 f 在点 x^0 处的 C -切微分
$\partial^M f(x^0)$	广义实值函数 f 在点 x^0 处的 M -次微分
$\partial^{**} f(x^0)$	广义实值函数 f 在点 x^0 处的 M -奇异次微分
$\psi'(x^0, y^0; d)$	集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 K -方向导数集合
$\psi'_w(x^0, y^0; d)$	集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 K -弱方向导数集合
$\psi_w^0(x^0, y^0; d)$	集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 GK -弱方向导数集合

$\phi_{\omega}^C(x^0, y^0; d)$	集值映射 ϕ 在点 (x^0, y^0) 处沿方向 d 的 CK-弱方向导数集合
$\phi^T(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的切导数
$\phi^D(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的 D -切导数
$\phi^C(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的 C -切导数
$\partial\phi(x^0, y^0)_p$	集值映射 ϕ 在点 (x^0, y^0) 处关于 p 的 K -次微分
$\partial_{\omega}\phi(x^0, y^0)_p$	集值映射 ϕ 在点 (x^0, y^0) 处关于 p 的 K -弱次微分
$\partial\phi(x^0)_p$	集值映射 ϕ 在点 x^0 处关于 p 的 K -次微分
$\partial_{\omega}\phi(x^0)_p$	集值映射 ϕ 在点 x^0 处关于 p 的 K -弱次微分
$\partial_{\omega}^0\phi(x^0, y^0)_p$	集值映射 ϕ 在点 (x^0, y^0) 处关于 p 的 GK-弱次微分
$\partial_{\omega}^C\phi(x^0, y^0)_p$	集值映射 ϕ 在点 (x^0, y^0) 处关于 p 的 CK-弱次微分
$\partial^T\phi(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的切微分
$\partial^C\phi(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的 C -切微分
$\partial\phi(x^0, y^0)_L$	集值映射 ϕ 在点 (x^0, y^0) 处的算子 K -次微分
$\partial_{\omega}\phi(x^0, y^0)_L$	集值映射 ϕ 在点 (x^0, y^0) 处的算子 K -弱次微分
$\partial_{\omega}^0\phi(x^0, y^0)_L$	集值映射 ϕ 在点 (x^0, y^0) 处的算子 GK-弱次微分
$\partial_{\omega}^C\phi(x^0, y^0)_L$	集值映射 ϕ 在点 (x^0, y^0) 处的算子 CK-弱次微分
$\partial^C\phi(x^0)_L$	集值映射 ϕ 在点 x^0 处的算子 CK-弱次微分
$\partial^M\phi(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的 M -次微分
$\partial^{MC}\phi(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的 MC -次微分
$\partial^{MF}\phi(x^0, y^0)$	集值映射 ϕ 在点 (x^0, y^0) 处的 MF -次微分

目 录

前 言 符号说明

第 1 章 凸集	1
§ 1.1 凸集及有关性质	1
§ 1.2 凸集的结构性质	11
§ 1.3 凸集分离定理	19
§ 1.4 凸锥和对偶锥	32
第 2 章 凸函数和凸映射	41
§ 2.1 凸函数及有关性质	41
§ 2.2 凸函数的连续性和对偶性	58
§ 2.3 广义凸函数	75
§ 2.4 凸映射和广义凸映射	88
第 3 章 凸函数的次微分	114
§ 3.1 凸函数的方向导数	114
§ 3.2 次梯度和次微分	120
§ 3.3 次微分的性质	130
§ 3.4 凸规划的最优性条件	142
第 4 章 Lipschitz 函数的 G -次微分	153
§ 4.1 G -方向导数	153
§ 4.2 G -次微分	156

§ 4.3	G -次微分的性质	168
§ 4.4	几何特性	183
§ 4.5	有限维情况	195
§ 4.6	Lipschitz 向量函数的 G -次微分	203
第 5 章	函数的切导数和切微分	209
§ 5.1	凸集的切锥	209
§ 5.2	D -切锥和 C -切锥	218
§ 5.3	D -切导数	224
§ 5.4	C -切导数和 C -切微分	228
§ 5.5	H -切导数	231
§ 5.6	若干关系	234
第 6 章	锥凸集值映射的锥次微分	244
§ 6.1	锥次微分和锥弱次微分	244
§ 6.2	存在性定理	249
§ 6.3	基本特性	266
§ 6.4	运算性质	275
§ 6.5	锥方向导数和锥弱方向导数	285
第 7 章	一般集值映射的锥弱次微分	294
§ 7.1	有效 Hahn-Banach 定理	294
§ 7.2	C -锥弱次微分	301
§ 7.3	G -锥弱次微分	309
§ 7.4	算子锥次微分	317
第 8 章	集值映射的切导数和切微分	323
§ 8.1	切导数和切微分	323
§ 8.2	D -切导数	332
§ 8.3	C -切导数和 C -切微分	335

§ 8.4 M -次微分	341
参考文献	361

第 1 章 凸 集

凸集是凸分析的研究对象和基本内容. 它的有关概念和理论, 则是凸分析和非光滑分析的基础和重要研究工具.

这一章将论述凸集及有关的概念和性质. 特别是, 介绍有着广泛应用的在不同空间情况的几个凸集分离定理. 同时还讨论作为重要特殊凸集的凸锥及其对偶锥.

§ 1.1 凸集及有关性质

本节阐述具有所谓凸性的集合及有关概念, 并且介绍它们的基本性质.

设 \mathcal{V} 是实线性空间, 它不止含有一个点.

定义 1.1.1 设集合 $S \subset \mathcal{V}$. 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda x^1 + (1 - \lambda)x^2 \in S \quad \forall x^1, x^2 \in S,$$

则称 S 是(\mathcal{V} 中的)凸集或集合 S 是凸的.

我们约定: 空集是凸集.

图 1.1.1 的(a)和(b)分别是凸集和不是凸集的图示.

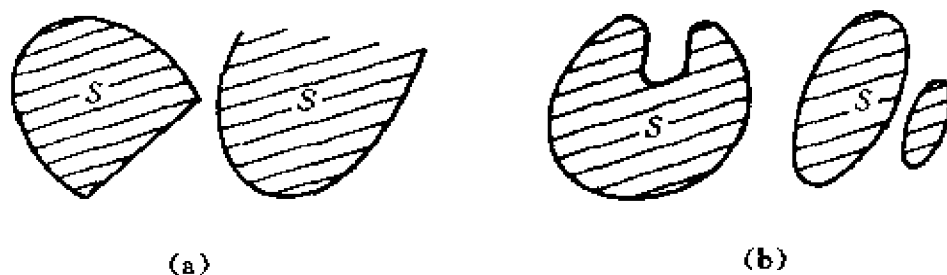


图 1.1.1

例 1.1.1 以下集合是相应空间中的凸集.

Euclid 空间 R^n 中的超平面:

$$H = \{x \in R^n \mid a^T x = \beta, a \in R^n, \beta \in R\}.$$

Euclid 空间 R^n 中的开半空间:

$$H_+^0 = \{x \in R^n \mid a^T x > \beta, a \in R^n, \beta \in R\}.$$

Banach 空间 \mathscr{B} 中以 x^0 为中心, δ 为半径的超球:

$$B_\delta(x^0) = \{x \in \mathscr{B} \mid \|x - x^0\| < \delta, x^0 \in \mathscr{B}, \delta > 0\}.$$

特别地, \mathscr{B} 中的开单位球:

$$B = \{x \in \mathscr{B} \mid \|x\| < 1\},$$

其中 $\|x\|$ 是 x 的范数.

线性空间 \mathscr{V} 中的闭线段:

$$[x^1, x^2] = \{x \in \mathscr{V} \mid x = \lambda x^1 + (1 - \lambda)x^2,$$

$$x^1, x^2 \in \mathscr{V}, \lambda \in [0, 1]\}.$$

线性空间 \mathscr{V} 中的正系数多项式集合:

$$P = \{p \in \mathscr{V} \mid p \text{ 是 } R \text{ 上正系数单变量多项式}\}.$$

先叙述关于凸集的一些简单性质.

定理 1.1.1 设集合 $S \subset \mathscr{V}$. S 是凸集当且仅当

$$(\alpha_1 + \alpha_2)S = \alpha_1 S + \alpha_2 S \quad \forall \alpha_1, \alpha_2 \in R_+. \quad (1.1.1)$$

证明 必要性. 当 $\alpha_1 = \alpha_2 = 0$ 时, 结论显然成立. 下设 $\alpha_1 + \alpha_2 > 0$.

先证等式(1.1.1)的右端包含左端. 为此, 设 $z \in (\alpha_1 + \alpha_2)S$, 则存在 $x \in S$ 使 $z = (\alpha_1 + \alpha_2)x = \alpha_1 x + \alpha_2 x$. 因为 $\alpha_1 x \in \alpha_1 S$, $\alpha_2 x \in \alpha_2 S$, 故 $z \in \alpha_1 S + \alpha_2 S$. 为证(1.1.1)的左端包含右端, 设 $z \in$

$\alpha_1 S + \alpha_2 S$, 则存在 $x^1, x^2 \in S$ 使

$$z = \alpha_1 x^1 + \alpha_2 x^2 = (\alpha_1 + \alpha_2) \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2} x^2 \right).$$

因为 S 是凸的, 由定义 1.1.1 有 $\frac{\alpha_1}{\alpha_1 + \alpha_2} x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2} x^2 \in S$, 于是得 $z \in (\alpha_1 + \alpha_2)S$.

充分性. 已知 (1.1.1) 成立. 对于任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 任取 $\alpha_1 > 0$, 令 $\alpha_2 = \frac{1-\lambda}{\lambda} \alpha_1$, 则有 $\lambda = \frac{\alpha_1}{\alpha_1 + \alpha_2}$, $1 - \lambda = \frac{\alpha_2}{\alpha_1 + \alpha_2}$. 由 (1.1.1) 可知 $\alpha_1 x^1 + \alpha_2 x^2 \in \alpha_1 S + \alpha_2 S = (\alpha_1 + \alpha_2)S$, 故存在 $x \in S$ 有 $\alpha_1 x^1 + \alpha_2 x^2 = (\alpha_1 + \alpha_2)x$ 或

$$\frac{\alpha_1}{\alpha_1 + \alpha_2} x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2} x^2 = x.$$

由此, 我们有 $\lambda x^1 + (1 - \lambda)x^2 = \frac{\alpha_1}{\alpha_1 + \alpha_2} x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2} x^2 = x \in S$, 于是由定义 1.1.1 得知 S 是凸的. \square

定理 1.1.2 设 $S_i \subset \mathcal{V} (i \in I, I \text{ 是指标集})$ 是凸集.

(1) 若 $I = \{1, \dots, m\}$ 是有限指标集, $\alpha_i \in R (i = 1, \dots, m)$, 则 $\sum_{i=1}^m \alpha_i S_i$ 是凸集.

(2) 若 I 是任意指标集, 则 $\bigcap_{i \in I} S_i$ 是凸集.

证明 由定义 1.1.1 直接可以推得. \square

定理 1.1.3 设 \mathcal{W} 是线性空间, $L: \mathcal{V} \rightarrow \mathcal{W}$ 是线性映射.

(1) 若 $S \subset \mathcal{V}$ 是凸集, 则 $L(S) \subset \mathcal{W}$ 是凸集.

(2) 若 $T \subset \mathcal{W}$ 是凸集, 则 $L^{-1}(T) \subset \mathcal{V}$ 是凸集.

证明 (1) 任取 $y^1, y^2 \in L(S)$, 则存在 $x^1, x^2 \in S$, 有 $y^1 = L(x^1)$, $y^2 = L(x^2)$. 对于任意的 $\lambda \in (0, 1)$, 因为 S 是凸的, 故 $\lambda x^1 + (1 - \lambda)x^2 \in S$. 于是, 由 L 是线性映射, 我们有

$$\lambda y^1 + (1 - \lambda)y^2 = \lambda L(x^1) + (1 - \lambda)L(x^2)$$

$$= L(\lambda x^1 + (1 - \lambda)x^2) \in L(S).$$

由于 y^1 和 y^2 是任取的, 根据定义 1.1.1 得知 $L(S)$ 是凸的.

(2) 由于 $L^{-1}: \mathscr{W} \rightarrow \mathscr{V}$ 也是线性映射, 与(1)的证明类似得证. \square

定义 1.1.2 设 $x^i \in \mathscr{V}$, $\lambda_i \geqslant 0$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda_i = 1$,

则称向量 $x = \sum_{i=1}^m \lambda_i x^i$ 是 x^1, \dots, x^m 的凸组合.

例 1.1.2 任一圆内的点都是圆周上某两点的凸组合. 任一三角形中的点都是它的三个顶点的凸组合.

定理 1.1.4 设集合 $S \subset \mathscr{V}$ 非空. S 是凸集当且仅当对任何正整数 $m \geqslant 2$ 有 $x^i \in S$, $\lambda_i \geqslant 0$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda_i = 1$, 使得

$$\sum_{i=1}^m \lambda_i x^i \in S. \quad (1.1.2)$$

证明 充分性. 在(1.1.2)中取 $m = 2$, 由定义 1.1.1 即知 S 是凸的.

必要性. 用数学归纳法证明: 当 $m = 2$ 时, 由 S 是凸的, 按定义 1.1.1 可知(1.1.2)成立.

设已知 $m = k$ 时(1.1.2)成立. 令

$$z = \sum_{i=1}^{k+1} \lambda_i x^i, \quad (1.1.3)$$

其中 $\lambda_i \geqslant 0$ ($i = 1, \dots, k+1$), $\sum_{i=1}^{k+1} \lambda_i = 1$. 若 $\lambda_{k+1} = 1$, 则 $z = x^{k+1} \in S$, 即(1.1.2)对于 $m = k+1$ 成立. 若 $\lambda_{k+1} \neq 1$, 因为(1.1.2)对于 $m = k$ 成立, 故有

$$z' = \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \in S,$$

$$\frac{\lambda_i}{1 - \lambda_{k+1}} \geq 0 \quad (i = 1, \dots, k), \quad \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} = 1.$$

由此, 根据(1.1.3)和 S 是凸的, 得到

$$z = \sum_{i=1}^k \lambda_i x^i + \lambda_{k+1} x^{k+1} = (1 - \lambda_{k+1}) z' + \lambda_{k+1} x^{k+1} \in S,$$

再由(1.1.3)便知(1.1.2)对于 $m = k + 1$ 成立. \square

定理 1.1.4 表明 S 是凸集的充分必要条件是: 该集合中任意个点的凸组合仍属于此集合.

以下两个概念是凸集的扩展.

定义 1.1.3 设集合 $S \subset \mathscr{V}$.

(1) 若对任意的 $\lambda \in R$ 有

$$\lambda x^1 + (1 - \lambda)x^2 \in S \quad \forall x^1, x^2 \in S,$$

则称 S 是(\mathscr{V} 中的)仿射集(仿射流形, 线性流形), 或集合 S 是仿射的.

(2) 若对任意的 $\lambda_1, \lambda_2 \in R$ 有

$$\lambda_1 x^1 + \lambda_2 x^2 \in S \quad \forall x^1, x^2 \in S,$$

则称 S 是(\mathscr{V} 中的)线性子空间.

我们约定: 空集是仿射集.

注 1.1.1 由定义 1.1.1 和定义 1.1.3 可知, 仿射集和线性子空间都是凸集. 又集合 S 是 \mathscr{V} 中的线性子空间, 即 S 对 \mathscr{V} 中的线性运算构成线性空间.

例 1.1.3 Euclid 空间 R^n 中的点、直线和超平面都是 R^n 中的仿射集. R^n 中过原点的直线和过原点的超平面都是 R^n 中的线性子空间.

定理 1.1.5 设集合 $S_i \subset \mathscr{V} (i \in I, I \text{ 是任意指标集})$.

(1) 若 $S_i (i \in I)$ 是仿射集, 则 $\bigcap_{i \in I} S_i$ 是仿射集.

(2) 若 $S_i (i \in I)$ 是线性子空间, 则 $\bigcap_{i \in I} S_i$ 是线性子空间.

证明 由定义 1.1.3 不难直接推得. \square

定义 1.1.4 设点 $x^i \in \mathcal{V}$ ($i = 1, \dots, m$).

(1) 若 $\lambda_i \in R$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda_i = 1$, 则称向量 $x = \sum_{i=1}^m \lambda_i x^i$ 是 x^1, \dots, x^m 的仿射组合.

(2) 若 $\lambda_i \in R$ ($i = 1, \dots, m$), 则称向量 $x = \sum_{i=1}^m \lambda_i x^i$ 是 x^1, \dots, x^m 的线性组合.

定理 1.1.6 设集合 $S \subset \mathcal{V}$ 非空.

(1) S 是仿射集当且仅当对任何正整数 $m \geq 2$, 有 $x^i \in S$, $\lambda_i \in R$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda_i = 1$, 使得 $\sum_{i=1}^m \lambda_i x^i \in S$.

(2) S 是线性子空间当且仅当对任何正整数 $m \geq 2$ 有 $x^i \in S$, $\lambda_i \in R$ ($i = 1, \dots, m$), 使得 $\sum_{i=1}^m \lambda_i x^i \in S$.

证明 与定理 1.1.4 的证明类似. \square

现在引进一般集合的凸包概念, 并讨论有关的一些性质.

定义 1.1.5 设集合 $S \subset \mathcal{V}$, 则 \mathcal{V} 中所有包含 S 的凸集的交称为 S 的凸包(或由 S 生成的凸集), 记作 $\text{co}S$.

我们约定: 空集的凸包是空集.

注 1.1.2 由定义 1.1.5, 集合的凸包是包含它的最小凸集(见图 1.1.2).

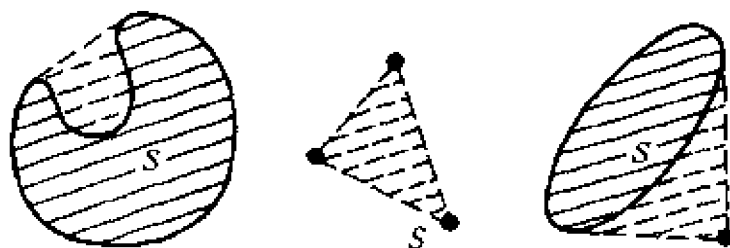


图 1.1.2

定理 1.1.7 设集合 $S \subset \mathcal{V}$.

(1) 若 S 是凸集, 则 $\text{co}S = S$.

(2) 若 $x \in \text{co}S$, 则 $\text{co}(S \cup \{x\}) = \text{co}S$.

证明 (1) 由定义 1.1.5 直接可知.

(2) 由 $x \in \text{co}S$, 因为 $S \subset \text{co}S$, 所以 $S \cup \{x\} \subset \text{co}S$. 由此, 我们有 $\text{co}(S \cup \{x\}) \subset \text{co}S$. 又显然有 $\text{co}S \subset \text{co}(S \cup \{x\})$, 故得结论. \square

下面是凸包的表示定理.

定理 1.1.8 设集合 $S \subset \mathcal{V}$, 则对任何正整数 m 有

$$\text{co}S = \left\{ x \in \mathcal{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S, \right. \\ \left. \forall \lambda_i \geq 0 \ (i = 1, \dots, m), \sum_{i=1}^m \lambda_i = 1 \right\}, \quad (1.1.4)$$

即 $\text{co}S$ 是 S 中向量的有限凸组合构成的集合.

证明 记 (1.1.4) 右端的集合为 D .

先证 $D \subset \text{co}S$. 当 $m=1$ 时, 显然成立. 对任何正整数 $m \geq 2$, 设 $x^1, \dots, x^m \in S$, 由定义 1.1.5, 有 $x^1, \dots, x^m \in \text{co}S$. 因为 $\text{co}S$ 是凸集, 由定理 1.1.4 知, 对任意的 $\lambda_i \geq 0$, $\sum_{i=1}^m \lambda_i = 1$, 有 $\sum_{i=1}^m \lambda_i x^i \in \text{co}S$, 所以 $D \subset \text{co}S$.

现证 $\text{co}S \subset D$. 任取 $x, y \in D$, 由 (1.1.4) 和 D 的定义可知, 存在正整数 m_1 和 m_2 以及 $x^i \in S$, $\lambda_i \geq 0$ ($i = 1, \dots, m_1$), $\sum_{i=1}^{m_1} \lambda_i = 1$; $y^j \in S$, $\mu_j \geq 0$ ($j = 1, \dots, m_2$), $\sum_{j=1}^{m_2} \mu_j = 1$, 使得 $x = \sum_{i=1}^{m_1} \lambda_i x^i$ 和 $y = \sum_{j=1}^{m_2} \mu_j y^j$. 由此, 对任意的 $\lambda \in (0, 1)$ 有

$$\lambda x + (1 - \lambda)y = \sum_{i=1}^{m_1} \lambda \lambda_i x^i + \sum_{j=1}^{m_2} (1 - \lambda) \mu_j y^j. \quad (1.1.5)$$

因为 $\lambda \lambda_i \geq 0$ ($i = 1, \dots, m_1$), $(1 - \lambda) \mu_j \geq 0$ ($j = 1, \dots, m_2$),

并且

$$\begin{aligned}\sum_{i=1}^{m_1} \lambda_i + \sum_{j=1}^{m_2} (1-\lambda) \mu_j &= \lambda \sum_{i=1}^{m_1} \lambda_i + (1-\lambda) \sum_{j=1}^{m_2} \mu_j \\ &= \lambda + (1-\lambda) = 1,\end{aligned}$$

所以由(1.1.5)和 D 的定义得到 $\lambda x + (1-\lambda)y \in D$. 据此, 由定义 1.1.1 即知 D 是凸集. 由于 $S \subset D$, 而 D 是凸集, 故由注 1.1.2 得到 $\text{co}S \subset D$. \square

定理 1.1.9 设集合 $S_1, S_2 \subset \mathcal{V}$.

(1) $\text{co}(S_1 \cap S_2) \subset \text{co}S_1 \cap \text{co}S_2$. 若 S_1 和 S_2 是凸集, 则 $\text{co}(S_1 \cap S_2) = \text{co}S_1 \cap \text{co}S_2$.

(2) $\text{co}(S_1 + S_2) = \text{co}S_1 + \text{co}S_2$.

(3) 若 S_1 和 S_2 是凸集, 则 $\text{co}(S_1 \cup S_2) = \bigcup_{\lambda \in [0, 1]} [\lambda S_1 + (1-\lambda)S_2]$.

证明 (1) 设 $y \in \text{co}(S_1 \cap S_2)$, 由定理 1.1.8 得知

$$\begin{aligned}y &\in \left\{ x \in \mathcal{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S_1 \cap S_2, \right. \\ &\quad \left. \forall \lambda_i \geq 0 (i = 1, \dots, m), \sum_{i=1}^m \lambda_i = 1 \right\}.\end{aligned}$$

由此, 我们有

$$\begin{aligned}y &\in \left\{ x \in \mathcal{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S_1, \right. \\ &\quad \left. \forall \lambda_i \geq 0 (i = 1, \dots, m), \sum_{i=1}^m \lambda_i = 1 \right\}\end{aligned}$$

和

$$\begin{aligned}y &\in \left\{ x \in \mathcal{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S_2, \right. \\ &\quad \left. \forall \lambda_i \geq 0 (i = 1, \dots, m), \sum_{i=1}^m \lambda_i = 1 \right\},\end{aligned}$$

从而 $y \in \text{co}S_1$ 和 $y \in \text{co}S_2$, 即 $y \in \text{co}S_1 \cap \text{co}S_2$.

设 S_1 和 S_2 是凸集, 我们再证明 $\text{co}S_1 \cap \text{co}S_2 \subset \text{co}(S_1 \cap S_2)$. 为此设 $y \in \text{co}S_1 \cap \text{co}S_2$, 因为 S_1 和 S_2 是凸的, 由定理 1.1.7 的 (1) 有 $y \in S_1 \cap S_2$. 据定义 1.1.5, 使得 $y \in S_1 \cap S_2 \subset \text{co}(S_1 \cap S_2)$.

(2) 先证 $\text{co}(S_1 + S_2) \subset \text{co}S_1 + \text{co}S_2$. 设 $y \in \text{co}(S_1 + S_2)$, 由定理 1.1.8 知 $y = \sum_{i=1}^m \lambda x^i$, 其中 $x^i \in S_1 + S_2$, $\lambda \geq 0$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda = 1$. 于是, 存在 $u^i \in S_1$, $v^i \in S_2$, 使 $x^i = u^i + v^i$ ($i = 1, \dots, m$) 有

$$y = \sum_{i=1}^m \lambda u^i + \sum_{i=1}^m \lambda v^i.$$

再由定理 1.1.8 得知 $y \in \text{co}S_1 + \text{co}S_2$.

为证 $\text{co}S_1 + \text{co}S_2 \subset \text{co}(S_1 + S_2)$, 设 $y \in \text{co}S_1 + \text{co}S_2$, 则存在 $u \in \text{co}S_1$ 和 $v \in \text{co}S_2$ 使 $y = u + v$. 由定理 1.1.8 我们有

$$u = \sum_{i=1}^{m_1} \lambda_i u^i, \text{ 其中 } u^i \in S_1, \lambda_i \geq 0 \ (i=1, \dots, m_1), \sum_{i=1}^{m_1} \lambda_i = 1,$$

$$v = \sum_{j=1}^{m_2} \mu_j v^j, \text{ 其中 } v^j \in S_2, \mu_j \geq 0 \ (j=1, \dots, m_2), \sum_{j=1}^{m_2} \mu_j = 1.$$

由此得到

$$y = \sum_{i=1}^{m_1} \lambda_i u^i + \sum_{j=1}^{m_2} \mu_j v^j = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \nu_{ij} (u^i + v^j),$$

其中 $u^i + v^j \in S_1 + S_2$, $\nu_{ij} = \lambda_i \mu_j \geq 0$ ($i = 1, \dots, m_1; j = 1, \dots, m_2$), $\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \nu_{ij} = 1$. 再由定理 1.1.8, 即得 $y \in \text{co}(S_1 + S_2)$.

(3) 设 $y \in \bigcup_{\lambda \in [0, 1]} [\lambda S_1 + (1 - \lambda) S_2]$, 则存在 $\bar{\lambda} \in [0, 1]$ 使

$y \in \bar{\lambda}S_1 + (1 - \bar{\lambda})S_2$. 于是, 存在 $x^1 \in S_1, x^2 \in S_2$, 有 $y \in \bar{\lambda}x^1 + (1 - \bar{\lambda})x^2$. 因为 $x^1, x^2 \in S_1 \cup S_2 \subset \text{co}(S_1 \cup S_2)$, 而 $\text{co}(S_1 \cup S_2)$ 是凸集, 故得 $y \in \text{co}(S_1 \cup S_2)$.

反之, 设 $y \in \text{co}(S_1 \cup S_2)$. 由定理 1.1.8 知, 存在正整数 m 和 $y^i \in S_1 \cup S_2, \bar{\lambda}_i \geq 0 (i = 1, \dots, m), \sum_{i=1}^m \bar{\lambda}_i = 1$, 使得

$$y = \sum_{i=1}^m \bar{\lambda}_i y^i. \quad (1.1.6)$$

若 $y^i \in S_1 (i = 1, \dots, m)$, 则因为 S_1 是凸的, 由 (1.1.6) 有 $y \in S_1$, 故 $y \in \bigcup_{\lambda \in [0, 1]} [\lambda S_1 + (1 - \lambda)S_2]$. 若 $y^i \in S_2 (i = 1, \dots, m)$, 则因为 S_2 是凸的, 同理有 $y \in \bigcup_{\lambda \in [0, 1]} [\lambda S_1 + (1 - \lambda)S_2]$. 现设 $y^i \in S_1 (i = 1, \dots, p)$ 和 $y^i \in S_2 (i = p + 1, \dots, m) (1 \leq p \leq m - 1)$.

令 $\bar{\lambda} = \sum_{i=1}^p \bar{\lambda}_i$, 则 $\bar{\lambda} \in (0, 1)$. 由 (1.1.6) 我们有

$$\begin{aligned} y &= \sum_{i=1}^p \bar{\lambda}_i y^i + \sum_{i=p+1}^m \bar{\lambda}_i y^i = \bar{\lambda} \sum_{i=1}^p \frac{\bar{\lambda}_i}{\bar{\lambda}} y^i + (1 - \bar{\lambda}) \sum_{i=p+1}^m \frac{\bar{\lambda}_i}{1 - \bar{\lambda}} y^i \\ &= \bar{\lambda} x^1 + (1 - \bar{\lambda}) x^2, \end{aligned} \quad (1.1.7)$$

其中 $x^1 = \sum_{i=1}^p \frac{\bar{\lambda}_i}{\bar{\lambda}} y^i, x^2 = \sum_{i=p+1}^m \frac{\bar{\lambda}_i}{1 - \bar{\lambda}} y^i$. 因为 S_1 和 S_2 是凸的, 故 $x^1 \in S_1$ 和 $x^2 \in S_2$. 据此, 由 (1.1.7) 得到

$$y = \bar{\lambda} x^1 + (1 - \bar{\lambda}) x^2 \in \bar{\lambda} S_1 + (1 - \bar{\lambda}) S_2,$$

从而 $y \in \bigcup_{\lambda \in [0, 1]} [\lambda S_1 + (1 - \lambda) S_2]$. \square

下面两个概念是凸包的扩展.

定义 1.1.6 设集合 $S \subset \mathcal{V}$.

(1) \mathcal{V} 中所有包含 S 的仿射集的交, 称为 S 的仿射包 (或由 S 生成的仿射集), 记作 $\text{aff} S$.

(2) \mathcal{V} 中所有包含 S 的线性子空间的交, 称为 S 的线性包

(或由 S 生成的线性子空间), 记作 $\text{lin}S$.

我们约定: 空集的仿射包是空集.

注 1.1.3 由定义 1.1.6 可知, 集合的仿射包是包含它的最小仿射集; 集合的线性包是包含它的最小线性子空间.

定理 1.1.10 设集合 $S \subset \mathscr{V}$.

(1) 若 S 是仿射集, 则 $\text{aff}S = S$. 若 $x \in \text{aff}S$, 则 $\text{aff}(S \cup \{x\}) = \text{aff}S$.

(2) 若 S 是线性子空间, 则 $\text{lin}S = S$. 若 $x \in \text{lin}S$, 则 $\text{lin}(S \cup \{x\}) = \text{lin}S$.

证明 与定理 1.1.7 的证明类似. \square

定理 1.1.11 设集合 $S \subset \mathscr{V}$.

(1) 对任何正整数 m 有

$$\begin{aligned} \text{aff}S = \left\{ x \in \mathscr{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S, \right. \\ \left. \forall \lambda_i \in R \ (i = 1, \dots, m), \sum_{i=1}^m \lambda_i = 1 \right\}. \end{aligned}$$

即 $\text{aff}S$ 是 S 中向量的有限仿射组合构成的集合.

(2) 对任何正整数 m 有

$$\begin{aligned} \text{lin}S = \left\{ x \in \mathscr{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S, \right. \\ \left. \forall \lambda_i \in R \ (i = 1, \dots, m) \right\}. \end{aligned}$$

即 $\text{lin}S$ 是 S 中向量的有限线性组合构成的集合.

证明 与定理 1.1.8 的证明类似. \square

§ 1.2 凸集的结构性质

本节介绍凸集的代数结构和拓扑结构性质. 先论述凸集的有关代数结构性质.

设 $x^1, x^2 \in \mathcal{V}$, 记 (x^1, x^2) 、 $[x^1, x^2]$ 以及 $(x^1, x^2]$ 和 $[x^1, x^2)$ 依次表示点 x^1 和 x^2 之间的开线段、闭线段以及半开半闭线段, 即

$$(x^1, x^2) = \{\lambda x^1 + (1 - \lambda)x^2 \mid x^1, x^2 \in \mathcal{V}, \lambda \in (0, 1)\},$$

$$[x^1, x^2] = \{\lambda x^1 + (1 - \lambda)x^2 \mid x^1, x^2 \in \mathcal{V}, \lambda \in [0, 1]\},$$

$$(x^1, x^2] = \{\lambda x^1 + (1 - \lambda)x^2 \mid x^1, x^2 \in \mathcal{V}, \lambda \in (0, 1]\},$$

$$[x^1, x^2) = \{\lambda x^1 + (1 - \lambda)x^2 \mid x^1, x^2 \in \mathcal{V}, \lambda \in [0, 1)\}.$$

(1.2.1)

定义 1.2.1 设集合 $S \subset \mathcal{V}$.

(1) 集合

$$S' = \{x \in S \mid \forall v \in \mathcal{V}, \exists \varepsilon > 0 \text{ 有 } [x, x + \varepsilon v] \subset S\}$$

称为 S 的代数内部. $x \in S'$ 称为 S 的代数内点. 若 $S = S'$, 则称 S 是代数开集.

(2) 集合

$$S'' = \{x \in S \mid \forall v \in \text{aff}S - x, \exists \varepsilon > 0 \text{ 有 } [x, x + \varepsilon v] \subset S\}$$

称为 S 的相对代数内部. $x \in S''$ 称为 S 的相对代数内点.

(3) 集合

$$S^s = \{x \in \mathcal{V} \mid \exists v \in S, \text{ 有 } [v, x) \subset S\}$$

称为 S 的代数闭包. 若 $S = S^s$, 则称 S 是代数闭集.

(4) 集合 $S^b = S^s \setminus S'$ 称为 S 的代数边界.

注 1.2.1 由定义 1.2.1, 显然有

$$S' \subset S'' \subset S \subset S^s.$$

定理 1.2.1 设 $S \subset \mathcal{V}$ 是凸集.

(1) 若 $x^1 \in S'$, $x^2 \in S^s$, 则 $[x^1, x^2) \subset S'$.

(2) 若 $x^1 \in S'' \neq \emptyset$, $x^2 \in S$, 则 $[x^1, x^2) \subset S''$.

证明 (1) 任取 $y \in [x^1, x^2)$, 则有 $\lambda \in [0, 1)$ 使 $y = \lambda x^2 +$

$(1 - \lambda)x^1$. 以下分两种情况讨论.

(a) 设 $x^2 \in S$. 因为 $x^1 \in S^i \subset S$, S 是凸的, 故 $y \in S$. 从 $x^1 \in S^i$, 由定义 1.2.1 的(1)可知, 对任意的 $v \in \mathcal{V}$, 存在 $\varepsilon > 0$ 有 $[x^1, x^1 + \varepsilon v] \subset S$, 从而 $x^1 + \varepsilon v \in S$. 由于 S 是凸的, 因此

$$y + \lambda \varepsilon v = \lambda x^2 + (1 - \lambda)(x^1 + \varepsilon v) \in S.$$

于是, 由 $y \in S$ 又知 $[y, y + \lambda \varepsilon v] \subset S$. 因为 $\lambda \varepsilon > 0$, 再按定义 1.2.1 的(1)有 $y \in S^i$, 由 y 的任意性得到 $[x^1, x^2) \subset S^i$.

(b) 设 $x^2 \in S^a \setminus S$. 这时 $x^2 \in S^a$, 由定义 1.2.1 的(3)可知存在 $v \in S$ 使 $[v, x^2) \subset S$. 若 x^1, x^2, v 三点共线, 不妨设 $v \in [x^1, x^2)$. 任取 $z \in [v, x^2) \subset S$, 则由(a)的结果有 $[x^1, z) \subset S^i$. 由 z 的任意性即得 $[x^1, x^2) \subset S^i$. 若 x^1, x^2, v 不共线, 则取 $t \in (x^2, y)$, 从而 $y \in (x^1, t)$. 从 $x^1 \in S^i$ 由定义 1.2.1 的(1)可知, 存在充分小的 $\varepsilon > 0$ 使 $u = x^1 + \varepsilon(x^2 - v) \in S$. 令 l 是过点 u 和 t 的直线, 由于 ε 充分小, 故 l 与 (v, x^2) 交于某点 $q \in S$. 据此, 由 S 是凸的, 得 $t \in S$. 于是, 由(a)也推知 $[x^1, t) \subset S^i$, 因而 $y \in S^i$. 由 y 的任意性即得 $[x^1, x^2) \subset S^i$.

(2) 任取 $y \in [x^1, x^2) \subset S$, 则有 $\lambda \in [0, 1)$ 使 $y = \lambda x^2 + (1 - \lambda)x^1$. 从 $x^1 \in S^i$ 由定义 1.2.1 中的(2)知, 对任意的 $v \in \text{aff} S - S$ 存在 $\varepsilon > 0$ 有 $[x^1, x^1 + \varepsilon v] \subset S$, 故 $x^1 + \varepsilon v \in S$. 因为 S 是凸的, $x^2 \in S$, 所以

$$y + (1 - \lambda)\varepsilon v = \lambda x^2 + (1 - \lambda)(x^1 + \varepsilon v) \in S.$$

从而 $[y, y + (1 - \lambda)\varepsilon v] \subset S$. 由此, 根据定义 1.2.1 的(2)知 $y \in S^i$. 由 y 的任意性即得结论. \square

定理 1.2.2 设 $S \subset \mathcal{V}$ 是凸集.

(1) S^i 是凸集.

(2) S^r 是凸集.

(3) S^a 是凸集.

证明 (1) 任取 $x^1, x^2 \in S^i \subset S$ 和 $\lambda \in (0, 1)$, 因为 S 是凸

的, 则 $y = \lambda x^1 + (1 - \lambda)x^2 \in S$. 由定义 1.2.1 的(1), 对任意的 $v \in \mathcal{V}$ 存在 $\varepsilon > 0$ 有

$$[x^1, x^1 + \varepsilon v] \subset S, [x^2, x^2 + \varepsilon v] \subset S,$$

从而 $x^1 + \varepsilon v, x^2 + \varepsilon v \in S$. 又因为 S 是凸的, 故

$$y + \varepsilon v = \lambda(x^1 + \varepsilon v) + (1 - \lambda)(x^2 + \varepsilon v) \in S.$$

据此, 由 $y \in S$ 知 $[y, y + \varepsilon v] \subset S$. 再按定义 1.2.1 的(1)得 $y \in S'$, 故 S' 是凸集.

(2) 与(1)的证明类似, 由定义 1.2.1 的(2)可得.

(3) 任取 $x^1, x^2 \in S^a$ 和 $\lambda \in (0, 1)$, 考虑 $y = \lambda x^1 + (1 - \lambda)x^2$. 由定义 1.2.1 的(3)可知存在 $v^1, v^2 \in S$ 使 $[v^1, x^1] \subset S, [v^2, x^2] \subset S$. 令 $v = \lambda v^1 + (1 - \lambda)v^2$, 则 $v \in S$. 再设 $\mu \in (0, 1)$, 则有

$$\mu v^1 + (1 - \mu)x^1 \in S, \mu v^2 + (1 - \mu)x^2 \in S.$$

因为

$$\begin{aligned} \mu v + (1 - \mu)y &= \mu[\lambda v^1 + (1 - \lambda)v^2] + (1 - \mu)[\lambda x^1 + (1 - \lambda)x^2] \\ &= \lambda[\mu v^1 + (1 - \mu)x^1] + (1 - \lambda)[\mu v^2 + (1 - \mu)x^2], \end{aligned}$$

所以 $\mu v + (1 - \mu)y \in S$, 或 $(v, y) \subset S$. 由此, 按定义 1.2.1 的(3)得到 $y \in S^a$, 故 S^a 是凸集. \square

定理 1.2.3 设 $T_1, T_2 \subset \mathcal{V}$ 是非空凸集, $T_1 \cap T_2 = \emptyset, T_1 \cup T_2 = \mathcal{V}$.

(1) $T_1^a \cap T_2^a = \emptyset, T_1^a \cap T_2^a = \emptyset$.

(2) $T_1^a \cup T_2^a = \mathcal{V}, T_1^a \cup T_2^a = \mathcal{V}$.

(3) 若 $H = T_1^a \cap T_2^a$, 则 H 是一仿射集.

(4) 若 $H = T_1^a \cap T_2^a, x^1 \in T_1, x^2 \in T_2$, 则存在 $z \in [x^1, x^2) \cap H$.

证明 (1) 先证明 $T_1^a \cap T_2^a = \emptyset$. 设若 $x \in T_2^a \setminus T_2$, 按定义 1.2.1 的(3)可知存在 $v \in T_2$, 使 $[v, x) \subset T_2$. 因此, 由已知

$T_1 \cap T_2 = \emptyset$ 有 $[v, x) \cap T_1 = \emptyset$, 所以 $x \notin T_1$. 设若 $x \in T_2$, 因为 T_2 是凸的, 则对任意的 $v \in T_2$ 都有 $[v, x) \subset T_2$, 故由 $T_1 \cap T_2 = \emptyset$ 得到 $x \notin T_1$. 将 T_1 和 T_2 对换, 即得 $T_1^* \cap T_2^* = \emptyset$.

(2) 先证 $T_1^* \cup T_2^* = \mathcal{X}$. 设 $x \notin T_2^*$, 由定义 1.2.1 的(3)和 $T_1 \cup T_2 = \mathcal{X}$, 则对任意的 $v \in \mathcal{X}$ 都存在 $y \in [v, x) \cap T_1$. 因为 T_1 是凸的, 于是 $[x, y] \subset T_1$. 由定义 1.2.1 的(1)即得 $x \in T_1^*$. 将 T_1 和 T_2 对换, 即得 $T_1^* \cup T_2^* = \mathcal{X}$.

(3) 由 T_1, T_2 是凸的, 根据定理 1.2.2 的(3)知 T_1^*, T_2^* 是凸的. 因此, 由定理 1.1.2 的(2)知 $H = T_1^* \cap T_2^*$ 是凸集, 故设 $x^1, x^2 \in H$, 则有 $[x^1, x^2] \subset H$. 为证 H 是仿射集, 反之假设存在 $\lambda \in R$, 使 $y = \lambda x^1 + (1 - \lambda)x^2 \notin H$. 因为 $[x^1, x^2] \subset H$, 则 $y \notin [x^1, x^2]$, 于是有 $x^1 \in (y, x^2)$ 或 $x^2 \in (x^1, y)$. 不妨设 $x^1 \in (y, x^2)$, 则由 $y \notin H$, 有 $y \in T_1^*$ 或 $y \in T_2^*$. 再不妨设 $y \in T_1^*$, 则由定理 1.2.1 的(1)得 $[y, x^2) \subset T_1^*$, 从而 $x^1 \in T_1^*$. 据此, 由(1)推得 $x^1 \notin T_2^*$, 导致与 $x^1 \in H$ 矛盾. 于是, 按定义 1.1.3 的(1)得知 H 是仿射集.

(4) 先设 $x^1 \notin T_1^*$. 由已证得的(2)推知 $x^1 \in T_2^*$, 再由 $x^1 \in T_1 \subset T_1^*$, 得 $x^1 \in T_1^* \cap T_2^* = H$. 因此, 存在 $z = x^1 \in [x^1, x^2) \cap H$. 再设 $x^1 \in T_1^*$. 反之假设 $[x^1, x^2) \cap H \neq \emptyset$, 则对任意的 $z \in [x^1, x^2)$, 有 $z \notin H = T_1^* \cap T_2^*$, 从而 $z \notin T_1^*$ 或 $z \notin T_2^*$. 不妨设 $z \notin T_1^*$, 则由(2)得 $z \in T_2^*$. 因为 $x^2 \in T_2^*$, T_2 是凸的, 故 $(x^2, z) \subset T_2$, 由 $z \in [x^1, x^2)$ 的任意性有 $(x^2, x^1) \subset T_2$, 于是 $[x^2, x^1) \subset T_2$. 据此, 由定义 1.2.1 的(3)得 $x^1 \in T_2^*$, 因而存在 $x^1 \in T_1^* \cap T_2^*$, 这导致与已知的(1)矛盾. \square

以下讨论凸集和凸包的拓扑结构性质. 设 \mathcal{X} 是线性拓扑空间.

定义 1.2.2 设集合 $S \subset \mathcal{X}$. 记点 $x \in \mathcal{X}$ 的邻域为 $U(x)$.

(1) S 的所有开子集的并称为 S 的内部, 记作 $\text{int}S$, 即

$$\text{int}S = \{x \in S \mid \exists U(x) \text{ 使 } U(x) \subset S\}.$$

$x \in \text{int}S$ 称为是 S 的内点. $S = \text{int}S$ 意味着 S 是开集.

(2) 集合

$$\text{ri}S = \{x \in S \mid \exists U(x) \text{ 使 } U(x) \cap \text{aff}S \subset S\}$$

称为 S 的相对内部. $x \in \text{ri}S$ 称为 S 的相对内点. 若 $S = \text{ri}S$, 则称 S 是相对开集.

(3) 所有包含 S 的闭集的交称为 S 的闭包, 记作 $\text{cl}S$ 或 \bar{S} . $x \in \text{cl}S$ 称为是 S 的接触点. $S = \text{cl}S$ 意味着 S 是闭集.

(4) 集合 $\partial S = \text{cl}S \setminus \text{int}S$ 称为 S 的边界. 集合 $\text{rb}S = \text{cl}S \setminus \text{ri}S$ 称为 S 的相对边界.

注 1.2.2 设 $S \subset \mathcal{H}$, 由定义 1.2.2, 显然有

$$\text{int}S \subset \text{ri}S \subset S \subset \text{cl}S.$$

又由定义 1.2.1 和定义 1.2.2, 可以推知

$$\text{int}S \subset S^{\circ}, S^{\circ} \subset \text{cl}S.$$

定理 1.2.4 设 $S \subset \mathcal{H}$ 是非空凸集, 则 $\text{ri}S = \text{int}S$ 当且仅当 $\text{int}S \neq \emptyset$.

证明 充分性. 设 $\text{int}S \neq \emptyset$, 则 S 是无限集, 从而 $\text{aff}S = \mathcal{H}$. 据此, 由定义 1.2.2 的(1)和(2), 我们有

$$\begin{aligned} \text{ri}S &= \{x \in S \mid \exists U(x) \text{ 使 } U(x) \cap \text{aff}S \subset S\} \\ &= \{x \in S \mid \exists U(x) \text{ 使 } U(x) \subset S\} = \text{int}S. \end{aligned}$$

必要性. 由 S 是凸的, 可以推知 $\text{ri}S \neq \emptyset$. 因此, $\text{int}S = \text{ri}S \neq \emptyset$. \square

定理 1.2.5 设 $S \subset \mathcal{H}$ 是非空凸集.

(1) 若 $x^1 \in \text{int}S$, $x^2 \in \text{cl}S$, 则 $[x^1, x^2) \subset \text{int}S$.

(2) 若 $x^1 \in \text{ri}S$, $x^2 \in \text{cl}S$, 则 $[x^1, x^2) \subset \text{ri}S$.

证明 (1) 任取 $y \in [x^1, x^2)$, 则有 $\lambda \in [0, 1)$ 使 $y = \lambda x^2 + (1 - \lambda)x^1$. 从 $x^1 \in \text{int}S$ 由定义 1.2.2 的(1)可知, 存在 x^1 的邻域 $U(x^1) \subset S$. 令 $W = \frac{1}{\lambda}[y - (1 - \lambda)U(x^1)]$, 则它是 x^2 的一个邻

域. 由 $x^2 \in \text{cl}S$ 知, 存在 $u^2 \in W \cap S$. 设 $u^2 = \frac{1}{\lambda}[y - (1 - \lambda)u^1]$, 其中 $u^1 \in U(x^1)$, 则 $y = \lambda u^2 + (1 - \lambda)u^1$. 因为 $U(x^1)$ 是 u^1 的邻域, 故 $U = \lambda u^2 + (1 - \lambda)U(x^1)$ 是 y 的邻域. 由于 $u^2 \in S$, $U(x^1) \subset S$, S 是凸集, 因此得知 $U \subset S$, 从而 $y \in \text{int}S$.

(2) 与 (1) 的证明类似, 由定义 1.2.2 的 (2) 可证明. \square

定理 1.2.6 设 $S \subset \mathcal{X}$ 是凸集.

(1) $\text{int}S$ 是凸集.

(2) $\text{ri}S$ 是凸集.

(3) $\text{cl}S$ 是凸集.

证明 (1) 若 $\text{int}S = \emptyset$, 则 $\text{int}S$ 是凸的. 设 $\text{int}S \neq \emptyset$, 任取 $x^1, x^2 \in \text{int}S$, 则 $x^1 \in \text{int}S \subset \text{cl}S$, $x^2 \in \text{int}S$. 由定理 1.2.5 的 (1), 有 $(x^1, x^2] \in \text{int}S$, 或按 (1.2.1) 知, 对任意的 $\lambda \in (0, 1)$ 有 $\lambda x^1 + (1 - \lambda)x^2 \in \text{int}S$. 据此, 由定义 1.1.1 便知 $\text{int}S$ 是凸的.

(2) 若 $\text{ri}S = \emptyset$, 则 $\text{ri}S$ 是凸的. 设 $\text{ri}S \neq \emptyset$, 任取 $x^1, x^2 \in \text{ri}S$, 则 $x^1 \in \text{ri}S \subset \text{cl}S$, $x^2 \in \text{ri}S$. 由定理 1.2.5 的 (2), 类似于 (1) 的证明可得证.

(3) 任取 $x^1, x^2 \in \text{cl}S$ 和任意的 $\lambda \in (0, 1)$, 考虑 $y = \lambda x^1 + (1 - \lambda)x^2$ 的邻域 $U(y)$. 由于集值映射 $U_y: \mathcal{X} \times \mathcal{X} \rightarrow 2^{\mathcal{X}}$, $(x^1, x^2) \mapsto \lambda x^1 + (1 - \lambda)x^2$ 是连续的, 故存在 x^1 的邻域 $U(x^1)$ 和 x^2 的邻域 $U(x^2)$, 使 $\lambda U(x^1) + (1 - \lambda)U(x^2) \subset U(y)$. 另外, 从 $x^1, x^2 \in \text{cl}S$ 知存在 $u^1 \in U(x^1) \cap S$, $u^2 \in U(x^2) \cap S$, 于是存在

$$\begin{aligned} \lambda u^1 + (1 - \lambda)u^2 &\in [\lambda U(x^1) + (1 - \lambda)U(x^2)] \cap S \\ &\subset U(y) \cap S. \end{aligned}$$

因为 $U(y) \cap S \neq \emptyset$, 所以 $y \in \text{cl}S$. \square

定理 1.2.7 设 $S \subset \mathcal{X}$ 是凸集.

(1) $\text{cl}S = \text{cl}(\text{cl}S) = \text{cl}(\text{ri}S)$.

(2) $\text{ri}S = \text{ri}(\text{cl}S) = \text{ri}(\text{ri}S)$.

(3) $\text{rb}S = \text{rb}(\text{cl}S) = \text{rb}(\text{ri}S)$.

(4) $\text{aff}S = \text{aff}(\text{cl}S) = \text{aff}(\text{ri}S)$.

证明 由定义 1.2.2, 利用定理 1.2.5 可以推证. \square

定理 1.2.8 设 $S \subset \mathcal{X}$ 是凸集.

(1) 若 $\text{int}S \neq \emptyset$, 则 $\text{int}S = S^\circ$, $\text{cl}S = S^{\circ\circ}$.

(2) 若 $\text{ri}S \neq \emptyset$, 则 $\text{ri}S = S^\circ$, $\text{cl}S = S^{\circ\circ}$.

证明 (1) 先证 $\text{int}S = S^\circ$. 由注 1.2.2 有 $\text{int}S \subset S^\circ$, 以下证 $S^\circ \subset \text{int}S$. 为此, 任取 $x \in S^\circ$, 设 $v \in \text{int}S$, 则存在 y 使 $x \in (y, v) \subset S$, 故 $y \in \text{cl}S$. 由定理 1.2.5 的(1)知 $(y, v) \subset \text{int}S$, 因为 $x \in (y, v)$, 于是得到 $x \in \text{int}S$.

以下证 $\text{cl}S = S^{\circ\circ}$. 由注 1.2.2, 有 $S^{\circ\circ} \subset \text{cl}S$. 为证 $\text{cl}S \subset S^{\circ\circ}$, 任取 $x \in \text{cl}S$, 设 $v \in \text{int}S$, 由定理 1.2.5 的(1)知 $[v, x) \subset \text{int}S$, 从而 $[v, x) \subset S$. 据此, 由定义 1.2.1 的(3)得 $x \in S^{\circ\circ}$.

(2) 与(1)的证明类似, 由定义 1.2.1 和定义 1.2.2, 利用定理 1.2.5 中的(2)可以推证. \square

定理 1.2.9 设 $S \subset \mathcal{X}$ 是开集, 则 $\text{co}S$ 是开集.

证明 由 S 是开集, 则 $S = \text{int}S$. 因为 $S \subset \text{co}S$, 故 $S = \text{int}S \subset \text{int}(\text{co}S)$. 由于 $\text{co}S$ 是凸的, 根据定理 1.2.6 的(1)知 $\text{int}(\text{co}S)$ 是凸集, 于是由注 1.1.2 有 $\text{co}S \subset \text{int}(\text{co}S)$. 又显然有 $\text{int}(\text{co}S) \subset \text{co}S$, 因此 $\text{co}S = \text{int}(\text{co}S)$, 故 $\text{co}S$ 是开集. \square

定义 1.2.3 设集合 $S \subset \mathcal{X}$, \mathcal{X} 中所有包含 S 的闭凸集的交称为是 S 的闭凸包, 记作 $\overline{\text{co}}S$.

注 1.2.3 由定义 1.2.3、定义 1.2.2 的(3)和定理 1.1.7 的(1), 若 S 是凸集, 则 $\overline{\text{co}}S = \overline{\text{co}}\overline{S} = \overline{S}$.

定理 1.2.10 设集合 $S \subset \mathcal{X}$ 非空, 则 $\overline{\text{co}}S = \text{cl}(\text{co}S)$.

证明 因为 $\text{co}S$ 是包含 S 的最小凸集, 而 $\overline{\text{co}}S$ 是包含 S 的闭凸集, 故 $\text{co}S \subset \overline{\text{co}}S$, 从而 $\text{cl}(\text{co}S) \subset \overline{\text{co}}S$. 反之, 由 $\overline{\text{co}}S$ 是包含 S 的最小闭凸集, 又有 $\overline{\text{co}}S \subset \text{cl}(\text{co}S)$. 而得知 $\overline{\text{co}}S = \text{cl}(\text{co}S)$. \square

注 1.2.4 闭凸包并不是闭集的凸包. 事实上, 闭集的凸包可以不是闭集. 而定理 1.2.10 表明, 闭凸包与凸包的闭包是相同的.

最后, 给出线性拓扑空间中严格凸集的概念.

定义 1.2.4 设集合 $S \subset \mathscr{E}$. 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda x^1 + (1 - \lambda)x^2 \in \text{int}S \quad \forall x^1, x^2 \in S, x^1 \neq x^2,$$

则称 S 是 (\mathscr{E}) 中的)严格凸集, 或集合 S 是严格凸的.

注 1.2.5 由定义 1.2.4 和定义 1.1.1 易知, 若 S 是严格凸集, 则 S 显然也是凸集, 但反之不然.

§ 1.3 凸集分离定理

这一节讨论在线性空间和线性拓扑空间中两个凸集被超平面分离的问题. 本节中介绍的凸集分离定理是凸分析和非光滑分析的重要分析工具.

先介绍几个有关的概念和引理.

设 \mathscr{V} 是线性空间, $\mathscr{L} \subset \mathscr{V}$ 是一线性子空间. 若存在 $v \in \mathscr{V} \setminus \mathscr{L}$, 使得

$$\mathscr{V} = \mathscr{L} + Rv = \{a + \lambda v \mid a \in \mathscr{L}, \lambda \in R\},$$

则称 \mathscr{L} 具有余维数 1.

定义 1.3.1 设 $\mathscr{L} \subset \mathscr{V}$ 是具有余维数 1 的线性子空间, $b \in \mathscr{V}$, 则称仿射集 $H = \mathscr{L} + b$ 是 \mathscr{V} 中的超平面.

设 \mathscr{V}^* 是 \mathscr{V} 的对偶空间, 即它是 \mathscr{V} 上线性泛函全体构成的空间. $\langle x^*, x \rangle$ 表示泛函 $x^* \in \mathscr{V}^*$ 在点 $x \in \mathscr{V}$ 处的值.

引理 1.3.1 设集合 $H \subset \mathscr{V}$ 非空, 则 H 是超平面当且仅当存在非零实线性泛函 $x^* \in \mathscr{V}^*$ 和实数 $\beta \in R$, 使得

$$H = \{x \in \mathscr{V} \mid \langle x^*, x \rangle = \beta\}. \quad (1.3.1)$$

证明 充分性. 设存在 $x^* \in \mathscr{V}^* \setminus \{0\}$ 和 $\beta \in R$ 使得 (1.3.1) 成立. 因为 x^* 是实线性泛函, 所以 $\mathscr{L} = \{x \in \mathscr{V} \mid \langle x^*, x \rangle = 0\}$ 是 \mathscr{V} 中的线性子空间. 又 x^* 是实线性泛函, 故存在 $v \in \mathscr{V}$ 使 $\langle x^*, v \rangle = 1$. 于是, 对任意的 $x \in \mathscr{V}$ 有 $\langle x^*, x - \langle x^*, x \rangle v \rangle = \langle x^*,$

$x\rangle - \langle x^*, x\rangle = 0$, 因而 $x - \langle x^*, x\rangle v \in \mathcal{L}$. 据此, 有 $x \in \mathcal{L} + \langle x^*, x\rangle v$, 故 $\mathcal{V} \subset \mathcal{L} + Rv$. 另外, 显然有 $\mathcal{L} + Rv \subset \mathcal{V}$, 因此得 $\mathcal{V} = \mathcal{L} + Rv$. 由于 \mathcal{L} 是具有余维数 1 的线性子空间, 所以由定义 1.3.1 知, $H = \mathcal{L} + \beta v$ 是 \mathcal{V} 中的超平面.

必要性. 设 $H = \mathcal{L} + b$ 是 \mathcal{V} 中的超平面, 其中 $\mathcal{L} \subset \mathcal{V}$ 是具有余维数 1 的线性子空间, $b \in \mathcal{V}$. 取 $v \in \mathcal{V} \setminus \mathcal{L}$, 则对任一 $x \in \mathcal{V}$ 有 $x = a + \lambda v$, 其中 $a \in \mathcal{L}$, $\lambda \in R$. 设 $x^* \in \mathcal{V}^* \setminus \{0\}$, 令 $\langle x^*, x\rangle = \lambda$, 并且 $\langle x^*, v\rangle = 1$, 当 $x \in \mathcal{L}$ 时 $\langle x^*, x\rangle = 0$. 再令 $\beta = \langle x^*, b\rangle$, 则显然有

$$H = \mathcal{L} + b \subset \{x \in \mathcal{V} \mid \langle x^*, x\rangle = \beta\}.$$

另外, 可以证明 $\{x \in \mathcal{V} \mid \langle x^*, x\rangle = \beta\} \subset H$, 因此得到 $H = \{x \in \mathcal{V} \mid \langle x^*, x\rangle = \beta\}$. \square

引理 1.3.2 设 $T_1, T_2 \subset \mathcal{V}$ 是非空凸集, $T_1 \cap T_2 = \emptyset$, $T_1 \cup T_2 = \mathcal{V}$. 若 $H = T_1^a \cap T_2^a$, 并且 $H \neq \mathcal{V}$, 则 H 是一超平面.

证明 设 $h \in H$, 因为 $H = (H - h) + h$, 由定理 1.2.3 的 (3) 知 $H = T_1^a \cap T_2^a$ 是仿射集. 因此, 只要证明线性子空间 $H - h$ 具有余维数 1, 由定义 1.3.1 即得 H 是一超平面.

设 $b \notin H = T_1^a \cap T_2^a$, 根据定理 1.2.3 的 (1) 可知, 若 $b \notin T_2^a$, 则 $b \in T_1^a$; 若 $b \notin T_1^a$, 则 $b \in T_2^a$. 于是有 $b \in T_1^a \cup T_2^a$. 不妨设 $b \in T_1^a$, 则有 $2h - b \notin H$ (否则, 将有 $b \in H$, 这与已设矛盾), 因此 $2h - b \in T_1^a \cup T_2^a$. 用反证法可得 $2h - b \notin T_1^a$ (因为若 $2h - b \in T_1^a$, 由 $b \in T_1^a$ 和 T_1^a 是凸的, 则有 $h \in T_1^a$, 导致矛盾), 从而知 $2h - b \in T_2^a$. 现在任取 $x \in \mathcal{V}$, 若 $x + h \in T_1^a$, 则 $[x + h, 2h - b] \cap H \neq \emptyset$, 故存在 $v \in H$ 和 $\lambda \in (0, 1]$ 使

$$v = \lambda(x + h) + (1 - \lambda)(2h - b),$$

由此, 我们有

$$x = \left(\frac{1}{\lambda} - 1\right)(b - h) + \frac{1}{\lambda}(v - h).$$

若 $x + h \in T_2$, 则 $[x + h, b) \cap H \neq \emptyset$, 故存在 $v \in H$ 和 $\lambda \in (0, 1]$ 使

$$v = \lambda(x + h) + (1 - \lambda)b,$$

从而

$$x = \left(1 - \frac{1}{\lambda}\right)(b - h) + \frac{1}{\lambda}(v - h).$$

综合以上讨论, 对任意的 $x \in \mathcal{V}$, 总存在 $y \in H - h$ 和 $\mu \in R$ 使得

$$x = y + \mu(b - h) \subset H - h + R(b - h).$$

据此, 得到 $\mathcal{V} = H - h + R(b - h)$, 这表明线性子空间 $H - h$ 具有余维数 1. 于是引理得证. \square

引理 1.3.3 设 $S_1, S_2 \subset \mathcal{V}$ 是凸集. 若 $S_1 \cap S_2 = \emptyset$, 则存在凸集 $T_1, T_2 \subset \mathcal{V}$, 使得 $S_1 \subset T_1, S_2 \subset T_2$, 并且 $T_1 \cap T_2 = \emptyset, T_1 \cup T_2 = \mathcal{V}$.

证明 设 $P, Q \subset \mathcal{V}$ 是凸集. 作集合

$$\mathcal{M} = \{(P, Q) \subset \mathcal{V} \times \mathcal{V} \mid S_1 \subset P, S_2 \subset Q, P \cap Q = \emptyset\},$$

并且定义 \mathcal{M} 中的序关系 \preceq 如下:

$$(P_1, Q_1) \preceq (P_2, Q_2) \iff P_1 \subset P_2, Q_1 \subset Q_2.$$

由于 \mathcal{M} 的每一全序子集 $\{(P_\alpha, Q_\alpha)\}$ 在 \mathcal{M} 中都有一个上界, 即 $\left(\bigcup_\alpha P_\alpha, \bigcup_\alpha Q_\alpha\right)$, 并且显然有 $S_1 \subset \bigcup_\alpha P_\alpha, S_2 \subset \bigcup_\alpha Q_\alpha$, 因而根据 Zorn 引理^[11], \mathcal{M} 必有极大元 (T_1, T_2) , 并且 $S_1 \subset T_1, S_2 \subset T_2, T_1 \cap T_2 = \emptyset$. 现在再证明有 $T_1 \cup T_2 = \mathcal{V}$. 事实上, 假若 $T_1 \cup T_2 \neq \mathcal{V}$, 则存在 $x \in \mathcal{V}$ 有 $x \notin T_1 \cup T_2$. 于是, 由 $T_1 \cap T_2 = \emptyset$ 可以推知 $\text{co}(\{x\} \cup T_1) \cap T_2 = \emptyset$ 或 $\text{co}(\{x\} \cup T_2) \cap T_1 = \emptyset$. 不妨设有 $\text{co}(\{x\} \cup T_1) \cap T_2 = \emptyset$, 记 $T'_1 = \text{co}(\{x\} \cup T_1)$, 则 T'_1 是凸集. 由 $S_1 \subset T_1 \subset T'_1$ 和 $S_2 \subset T_2$ 得知 $(T'_1, T_2) \in \mathcal{M}$, 这表明 (T_1, T_2) 不是极大元, 导致与已知矛盾. \square

设有超平面 (1.3.1), 我们称集合

$$H^{\circ} = \{x \in \mathcal{V} \mid \langle x^*, x \rangle \leq \beta\}$$

和

$$H^{\circ}_+ = \{x \in \mathcal{V} \mid \langle x^*, x \rangle \geq \beta\}$$

是由超平面 H 确定的闭半空间; 称集合

$$H^{\circ}_- = \{x \in \mathcal{V} \mid \langle x^*, x \rangle < \beta\}$$

和

$$H^{\circ}_+ = \{x \in \mathcal{V} \mid \langle x^*, x \rangle > \beta\}$$

是由超平面 H 确定的开半空间. 显然, 上述 4 个半空间构成两组互补的凸集, 即

$$H^{\circ} \cup H^{\circ}_+ = \mathcal{V}, \quad H^{\circ} \cap H^{\circ}_+ = \emptyset,$$

$$H^{\circ}_- \cup H^{\circ}_+ = \mathcal{V}, \quad H^{\circ}_- \cap H^{\circ}_+ = \emptyset.$$

下面讨论两个集合被超平面分离的问题.

定义 1.3.2 设集合 $S_1, S_2 \subset \mathcal{V}$ 非空.

(1) 若存在超平面 $H \subset \mathcal{V}$ 使 $S_1 \subset H^{\circ}_-$ 和 $S_2 \subset H^{\circ}_+$, 即存在 $x^* \in \mathcal{V}^* \setminus \{0\}$ 和 $\beta \in R$ 使

$$\langle x^*, x^1 \rangle \leq \beta \leq \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2, \quad (1.3.2)$$

则称集合 S_1 和 S_2 被超平面 H 分离, 或 H 分离 S_1 和 S_2 . 若再有 $S_1 \cup S_2 \not\subset H$, 则称 H 正常分离 S_1 和 S_2 , 否则, 称 H 非正常分离 S_1 和 S_2 .

(2) 若存在超平面 $H \subset \mathcal{V}$ 使 $S_1 \subset H^{\circ}_-$ 和 $S_2 \subset H^{\circ}_+$ 或 $S_1 \subset H^{\circ}_+$ 和 $S_2 \subset H^{\circ}_-$, 即存在 $x^* \in \mathcal{V}^* \setminus \{0\}$ 和 $\beta \in R$ 使

$$\langle x^*, x^1 \rangle \leq \beta < \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2, \quad (1.3.3)$$

或

$$\langle x^*, x^1 \rangle < \beta \leq \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2. \quad (1.3.4)$$

则称集合 S_1 和 S_2 被超平面 H 严格分离, 或 H 严格分离 S_1 和 S_2 .

(3) 若存在 $x^* \in \mathcal{V}^* \setminus \{0\}$ 和 $\beta \in R$, 使

$$\sup_{x^1 \in S_1} \langle x^*, x^1 \rangle \leq (<) \beta < (\leq) \inf_{x^2 \in S_2} \langle x^*, x^2 \rangle, \quad (1.3.5)$$

则称集合 S_1 和 S_2 被超平面 H 强分离, 或 H 强分离 S_1 和 S_2 .

注 1.3.1 显然, 若超平面 H 强分离集合 S_1 和 S_2 , 则它必严格分离 S_1 和 S_2 ; 若 H 严格分离 S_1 和 S_2 , 则它分离 S_1 和 S_2 ; 反之不然.

图 1.3.1 依次是正常分离, 非正常分离, 严格分离和强分离的情况.

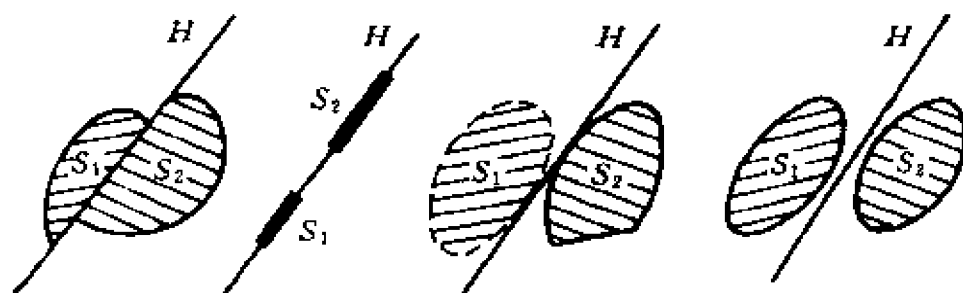


图 1.3.1

下面我们给出线性空间 \mathcal{V} 中的几个凸集分离定理.

定理 1.3.4 设 $S_1, S_2 \subset \mathcal{V}$ 是凸集, $S_1 \neq \emptyset, S_2 \neq \emptyset$, 并且 $S_1 \cap S_2 = \emptyset$.

(1) 存在超平面 $H \subset \mathcal{V}$ 正常分离 S_1 和 S_2 .

(2) 存在超平面 $H \subset \mathcal{V}$ 严格分离 S_1 和 S_2 .

证明 (1) 因为 S_2 是凸的, 由定理 1.2.2 的(1), 则 S_2 是凸的. 又由 S_1 是凸的和 $S_1 \cap S_2 = \emptyset$, 根据引理 1.3.3 可知, 存在凸集 $T_1, T_2 \subset \mathcal{V}$ 使 $S_1 \subset T_1, S_2 \subset T_2$, 并且 $T_1 \cap T_2 = \emptyset, T_1 \cup T_2 = \mathcal{V}$. 从 S_1 和 S_2 是非空凸集可知 T_1 和 T_2 是非空凸集, 于是由

定理 1.2.3 的(1)得

$$T_1^a \cap T_2^a = \emptyset. \quad (1.3.6)$$

另外,由 $S_2^i \subset T_2$ 有 $(S_2^i)^i \subset T_2^i$, 从定义 1.2.1 的(1)和定理 1.2.1 的(1)可推知 $(S_2^i)^i = S_2^i$, 故 $S_2^i \subset T_2^i$. 因此,由(1.3.6)我们有

$$T_1^a \cap S_2^i = \emptyset. \quad (1.3.7)$$

设 $H = T_1^a \cap T_2^a$, 因为 $S_2^i \neq \emptyset$, 故存在 $y \in S_2^i \subset \mathcal{V}$, 由(1.3.7)知 $y \notin T_1^a$. 于是有 $y \notin T_1^a \cap T_2^a = H$, 从而 $H \neq \mathcal{V}$. 据此,由引理 1.3.2 得知 H 是一超平面. 再由引理 1.3.1, 可设

$$H = \{x \in \mathcal{V} \mid \langle x^*, x \rangle = \beta\}, \quad (1.3.8)$$

其中 $x^* \in \mathcal{V}^* \setminus \{0\}$, $\beta \in R$. 因为据(1.3.7)有

$$H \cap S_2^i = T_1^a \cap T_2^a \cap S_2^i = \emptyset, \quad (1.3.9)$$

所以对任意的 $x \in S_2^i$, 由(1.3.8)有 $\langle x^*, x \rangle \neq \beta$. 不妨设存在 $v \in S_2^i$ 使 $\langle x^*, v \rangle > \beta$. 我们证明

$$\langle x^*, x^2 \rangle \geq \beta \quad \forall x^2 \in S_2. \quad (1.3.10)$$

事实上,反之,若假设存在某 $\bar{x}^2 \in S_2$, 使 $\langle x^*, \bar{x}^2 \rangle < \beta$, 则因 x^* 是线性泛函,故存在 $z \in (v, \bar{x}^2)$, 使 $\langle x^*, z \rangle = \beta$, 从而 $z \in H$. 另外,从 $v \in S_2^i$ 和 $\bar{x}^2 \in S_2 \subset S_2^i$ 由定理 1.2.1 的(1)有 $[v, \bar{x}^2] \subset S_2^i$, 故 $z \in S_2^i$. 于是,再由 $z \in H$ 得知存在 $z \in H \cap S_2^i$, 这导致与(1.3.9)矛盾. 以下再证明有

$$\langle x^*, x^1 \rangle \leq \beta \quad \forall x^1 \in S_1. \quad (1.3.11)$$

仍用反证法,假设存在 $\bar{x}^1 \in S_1$ 使 $\langle x^*, \bar{x}^1 \rangle > \beta$. 从 $\bar{x}^1 \in S_1 \subset T_1$ 和 $v \in S_2^i \subset T_2^i$, 由定理 1.2.3 的(4)可知存在 $z' \in [\bar{x}^1, v] \cap H$, 依据(1.3.8)即有 $z' \in [\bar{x}^1, v]$ 使 $\langle x^*, z' \rangle = \beta$. 但已知 $\langle x^*, \bar{x}^1 \rangle > \beta$ 和 $\langle x^*, v \rangle > \beta$, 并且 $z' \in [\bar{x}^1, v]$, 因此 $\langle x^*, z' \rangle > \beta$, 而这与 $\langle x^*, z' \rangle = \beta$ 相矛盾. 由(1.3.10)和(1.3.11)得(1.3.2), 故

按定义 1.3.2 的(1)知 H 分离 S_1 和 S_2 . 又由(1.3.9)知 $S_2 \not\subset H$, 所以 H 正常分离 S_1 和 S_2 .

(2) 从(1.3.9)和(1.3.8)可知

$$\langle x^*, x \rangle \neq \beta \quad \forall x \in S_2'.$$

再由(1.3.10)我们有

$$\langle x^*, x^2 \rangle > \beta \quad \forall x^2 \in S_2'.$$

于是,注意到(1.3.11),由定义 1.3.2 的(2)即知 H 严格分离 S_1 和 S_2' . \square

定理 1.3.5 设 $S_1, S_2 \subset \mathcal{V}$ 是凸集, $S_1'' \neq \emptyset, S_2'' \neq \emptyset$, 并且 $S_1'' \cap S_2'' = \emptyset$.

(1) 存在超平面 $H \subset \mathcal{V}$ 分离 S_1 和 S_2 .

(2) 存在超平面 $H \subset \mathcal{V}$ 严格分离 S_1 和 S_2'' 或 S_1'' 和 S_2 .

证明 令 $S = S_2 - S_1$, 因为 S_1 和 S_2 是凸的, 由定理 1.1.2 中的(1)知 S 是凸的. 据此, 由定义 1.2.1 的(2)和定理 1.2.1 的(2), 可以推知

$$S'' = (S_2 - S_1)'' = S_2'' - S_1''.$$

于是, 由 $S_1'' \cap S_2'' = \emptyset$ 得

$$0 \notin S''. \quad (1.3.12)$$

(1) 设 $0 \notin \text{aff} S$. 由定义 1.1.6 和定义 1.1.3 可知, 存在 $\bar{z} \in S$ 和 $z_0 \notin \text{lin}(S - \bar{z})$ 使

$$\text{aff} S = z_0 + \text{lin}(S - \bar{z}). \quad (1.3.13)$$

因为 $\text{lin}(S - \bar{z}) \subset \mathcal{V}$ 是维数不为 0 的真线性子空间, 设 $\{e_i\}_{i \in I}$ 是 $\text{lin}(S - \bar{z})$ 的 Hamel 基^[18], 则 $\{z_0\} \cup \{e_i\}_{i \in I}$ 可扩充为 \mathcal{V} 的基, 记作 $\{z_0\} \cup \{e_j\}_{j \in J} (I \subset J)$. 由于任何 $z^* \in \mathcal{V}^* \setminus \{0\}$ 都可由实数组 $\{\langle z^*, z_0 \rangle\} \cup \{\langle z^*, e_j \rangle\}_{j \in J}$ 唯一确定, 所以存在 $x^* \in \mathcal{V}^* \setminus \{0\}$, 使

$$\langle x^*, z_0 \rangle = 1 \quad (1.3.14)$$

和

$$\langle x^*, e_i \rangle = 0, i \in I \subset J. \quad (1.3.15)$$

又因为任意的 $z \in \text{lin}(S - \bar{z})$ 是 $\{e_i\}_{i \in I}$ 的线性组合, 故由(1.3.15)得

$$\langle x^*, z \rangle = 0 \quad \forall z \in \text{lin}(S - \bar{z}). \quad (1.3.16)$$

从(1.3.13)可知, 对任意的 $x \in S \subset \text{aff}S$, 存在 $z' \in \text{lin}(S - \bar{z})$ 使 $x = z_0 + z'$, 故由(1.3.14)和(1.3.16)知

$$\langle x^*, x \rangle = \langle x^*, z_0 \rangle + \langle x^*, z' \rangle = 1.$$

因为 $S = S_2 - S_1$, 令 $x = x^2 - x^1$, $x^1 \in S_1$, $x^2 \in S_2$, 由 $x \in S$ 的任意性, 从上式有 $\langle x^*, x^2 - x^1 \rangle = 1$, 或

$$\langle x^*, x^1 \rangle = \langle x^*, x^2 \rangle - 1 \quad \forall x^1 \in S_1, \forall x^2 \in S_2.$$

令 $\beta = \sup_{x^1 \in S_1} \langle x^*, x^1 \rangle + \frac{1}{2}$, 则得到

$$\langle x^*, x^1 \rangle < \beta < \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2. \quad (1.3.17)$$

由此, 根据定义 1.3.2 的(1), 结论成立.

设 $\theta \in \text{aff}S$. 这时 $\mathcal{V}_\theta = \text{aff}S = \text{lin}S$ 是 \mathcal{V} 的线性子空间, 并且由定义 1.2.1 的(1)和(2)知 S^π 即 S 在 \mathcal{V}_θ 上的代数内部. 因此, 由(1.3.12)和注 1.2.1, 利用定理 1.3.4 的(1)得知, 存在超平面分离 $\{\theta\}$ 和 S , 即存在 $x_1^* \in \mathcal{V}_\theta^* \setminus \{\theta\}$ 和 $\beta_1 \in R$ 使得

$$0 \leq \beta_1 \leq \langle x_1^*, y \rangle \quad \forall y \in S.$$

将 x_1^* 延拓到整个空间, 则存在 $x^* \in \mathcal{V}^* \setminus \{\theta\}$ 有

$$0 \leq \langle x_1^*, y \rangle = \langle x^*, y \rangle \quad \forall y \in S.$$

令 $y = x^2 - x^1$, $x^1 \in S_1$, $x^2 \in S_2$, 从上式得到

$$\langle x^*, x^1 \rangle \leq \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2.$$

取 $\beta = \sup_{x^1 \in S_1} \langle x^*, x^1 \rangle$, 由定义 1.3.2 的(1)即得知 H 分离 S_1 和 S_2 .

(2) 设 $\theta \notin \text{aff} S$. 由(1.3.17)和 $S_1^r \subset S_1, S_2^r \subset S_2$, 按定义 1.3.2 中的(2)知结论成立.

设 $\theta \in \text{aff} S$. 与(1)中的证明相同, 由(1.3.12)利用定理 1.3.4 中的(2)得知, 存在 $x_2^* \in \mathcal{V}_\theta^* \setminus \{\theta\}$ 和 $\beta_2 \in R$ 使得

$$0 \leq \beta_2 < \langle x_2^*, y \rangle \quad \forall y \in S^r.$$

将 x_2^* 延拓到整个空间, 则存在 $x^* \in \mathcal{V}^* \setminus \{\theta\}$ 有

$$0 \leq \beta_2 < \langle x_2^*, y \rangle = \langle x^*, y \rangle \quad \forall y \in S^r.$$

令 $y = x^2 - x^1, x^1 \in S_1^r, x^2 \in S_2^r$, 从 $S^r = S_2^r - S_1^r$ 和上式, 有

$$\langle x^*, x^1 \rangle < \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1^r, \forall x^2 \in S_2^r. \quad (1.3.18)$$

取 $\beta' = \sup_{x^1 \in S_1} \langle x^*, x^1 \rangle, \beta'' = \inf_{x^2 \in S_2} \langle x^*, x^2 \rangle$, 则有

$$\langle x^*, x^1 \rangle \leq \beta' \leq \beta'' \leq \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1^r, \forall x^2 \in S_2.$$

若 $\beta' < \beta''$, 则

$$\langle x^*, x^1 \rangle \leq \beta' < \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1^r, \forall x^2 \in S_2. \quad (1.3.19)$$

若 $\beta' = \beta''$, 则不存在 $\bar{x}^1 \in S_1^r$ 和 $\bar{x}^2 \in S_2^r$ 使

$$\langle x^*, \bar{x}^1 \rangle = \beta' = \beta'' = \langle x^*, \bar{x}^2 \rangle,$$

否则, 与(1.3.18)相矛盾. 因此得知或者有(1.3.19), 或者有

$$\langle x^*, x^1 \rangle < \beta'' \leq \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2^r.$$

于是, 由定义 1.3.2 的(2)即得 H 严格分离 S_1 和 S_2^r 或 S_1^r 和 S_2 . [

注 1.3.2 因为 $S_1^r \subset S_1, S_2^r \subset S_2$, 所以在定理 1.3.5 的条件下, 由定理 1.3.5 的(2)还可推知存在超平面 $H \subset \mathcal{V}$ 严格分离 S_1^r

和 S_2° .

定理 1.3.6 设集合 $S_1, S_2 \subset \mathscr{V}$ 非空. 若 $S_1 - S_2$ 是凸集, $(S_1 - S_2)^{\circ} \neq \emptyset$, 并且 $\theta \notin (S_1 - S_2)^{\circ}$, 则存在超平面 $H \subset \mathscr{V}$ 强分离 S_1 和 S_2 .

证明 令 $S = S_1 - S_2$, 则 $S^{\circ} \neq \emptyset$, $\theta \notin S^{\circ}$.

设 $\theta \notin \text{aff} S$. 由定理 1.3.5 的证明可知, 存在 $x^* \in \mathscr{V}^* \setminus \{\theta\}$ 使

$$\langle x^*, x^1 \rangle = \langle x^*, x^2 \rangle - 1 \quad \forall x^1 \in S_1, \forall x^2 \in S_2.$$

由此, 按定义 1.3.2 的(3)可以推知结论成立.

设 $\theta \in \text{aff} S$. 取 $v \in S^{\circ}$, 因为 $\theta \notin S^{\circ}$, 则存在 $\bar{v} \in (\theta, v)$ 使 $[\bar{v}, v) \subset S^{\circ}$, $[\theta, \bar{v}) \notin S^{\circ}$, 故有 $\bar{v} \notin S^{\circ}$. 据此, 视 $\{\bar{v}\}$ 为单点凸集, 又因 S 是凸的, 由定理 1.2.2 的(2)知 S° 是凸的, 于是由定理 1.3.5 的(2)知存在 $x^* \in \mathscr{V}^* \setminus \{\theta\}$ 使

$$\langle x^*, v \rangle < \langle x^*, \bar{v} \rangle \quad \forall v \in S^{\circ},$$

从而得

$$\sup_{v \in S} \langle x^*, v \rangle \leqslant \langle x^*, \bar{v} \rangle. \quad (1.3.20)$$

由于 $\bar{v} \in (\theta, v)$, 则存在 $\lambda \in (0, 1)$ 使 $\bar{v} = \lambda v$, 又由 $v \in S^{\circ}$, 故

$$\langle x^*, v \rangle < \langle x^*, \bar{v} \rangle = \lambda \langle x^*, v \rangle,$$

从而有 $\langle x^*, v \rangle < 0$ 和 $\langle x^*, \bar{v} \rangle < 0$. 根据 (1.3.20), 则得 $\sup_{v \in S} \langle x^*, v \rangle < 0$. 由 $S = S_1 - S_2$, 令 $v = x^1 - x^2$, $x^1 \in S_1$, $x^2 \in S_2$, 我们有

$$\sup_{\substack{x^1 \in S_1 \\ x^2 \in S_2}} \langle x^*, x^1 - x^2 \rangle < 0,$$

因而得到

$$\sup_{x^1 \in S_1} \langle x^*, x^1 \rangle < \inf_{x^2 \in S_2} \langle x^*, x^2 \rangle.$$

据此, 由定义 1.3.2 中的(3)得知结论成立. \square

以下介绍线性拓扑空间 \mathscr{X} 中的凸集分离定理.

设 \mathcal{X}^* 是 \mathcal{X} 的对偶空间, 即 \mathcal{X}^* 是 \mathcal{X} 上连续线性泛函全体构成的空间. 记 $\langle x^*, x \rangle$ 表示连续线性泛函 $x^* \in \mathcal{X}^*$ 在点 $x \in \mathcal{X}$ 处的值. 可以证明, 当 $x^* \in \mathcal{X}^* \setminus \{0\}$ 和 $\beta \in \mathbb{R}$ 时, \mathcal{X} 中由连续线性泛函确定的超平面

$$H = \{x \in \mathcal{X} \mid \langle x^*, x \rangle = \beta\}$$

是闭集^[14]. 这时, H 是 \mathcal{X} 中的闭超平面.

定理 1.3.7 设 $S_1, S_2 \subset \mathcal{X}$ 是凸集, $S_1 \neq \emptyset$, $\text{int}S_2 \neq \emptyset$, 并且 $S_1 \cap \text{int}S_2 = \emptyset$.

- (1) 存在闭超平面 $H \subset \mathcal{X}$ 正常分离 S_1 和 S_2 .
- (2) 存在闭超平面 $H \subset \mathcal{X}$ 严格分离 S_1 和 $\text{int}S_2$.

证明 由 $\text{int}S_2 \neq \emptyset$ 和定理 1.2.8 的(1)有 $\text{int}S_2 = S_2'$. 注意到由 \mathcal{X} 上的连续线性泛函确定的超平面是闭的, 利用定理 1.3.4 即可得到结论. \square

定理 1.3.8 设 \mathcal{X} 是局部凸线性拓扑空间, $S_1, S_2 \subset \mathcal{X}$ 是凸集, $\text{ri}S_1 \neq \emptyset$, $\text{ri}S_2 \neq \emptyset$, 并且 $\text{ri}S_1 \cap \text{ri}S_2 = \emptyset$.

- (1) 存在闭超平面 $H \subset \mathcal{X}$ 分离 S_1 和 S_2 .
- (2) 存在闭超平面 $H \subset \mathcal{X}$ 严格分离 $\text{ri}S_1$ 和 $\text{ri}S_2$.

证明 由 $\text{ri}S_1 \neq \emptyset$ 和 $\text{ri}S_2 \neq \emptyset$, 根据定理 1.2.8 的(2), 我们有 $\text{ri}S_1 = S_1''$ 和 $\text{ri}S_2 = S_2''$. 于是, 由定理 1.3.5 和注 1.3.2, 注意到这时超平面是闭的即得结论. \square

定理 1.3.9 设 \mathcal{X} 是分离局部凸线性拓扑空间, $S_1 \subset \mathcal{X}$ 是紧凸集, $S_2 \subset \mathcal{X}$ 是闭凸集. 若 $S_1 \cap S_2 = \emptyset$, 则存在闭超平面强分离 S_1 和 S_2 .

证明 因为 $S_1 - S_2$ 是闭凸集, 而 $0 \notin S_1 - S_2$, 从而存在凸零邻域 U , 使得 $U \cap (S_1 - S_2) = \emptyset$. 由此, 根据文献[25]的定理 3.2.10 即可得证. \square

定理 1.3.10 设 $S_1, S_2 \subset \mathcal{X}$ 是开凸集, $S_1 \neq \emptyset$, $S_2 \neq \emptyset$. 若 $S_1 \cap S_2 = \emptyset$, 则存在闭超平面严格分离 S_1 和 S_2 .

证明 因 S_2 是开凸集, 故 $\text{int}S_2 = S_2$. 由 $S_1 \neq \emptyset$, $\text{int}S_2 = S_2$

$\neq \emptyset$ 和 $S_1 \cap \text{int}S_2 = S_1 \cap S_2 = \emptyset$, 根据定理 1.3.7 的(2)可得结论. \square

最后, 利用凸集分离定理, 阐述凸集在一点处被超平面支撑以及闭凸集由它的支撑函数表示的问题.

定义 1.3.3 设集合 $S \subset \mathcal{V}$ 非空, 点 $x^0 \in S^b$, $H \subset \mathcal{V}$ 是超平面. 若 $x^0 \in H$, 并且 $S \subset H_+^c$ 或 H_-^c , 则称 H 是集合 S 在点 x^0 处的支撑超平面, 或称 H 在点 x^0 处支撑 S , x^0 称为支撑点. 再若有 $S \not\subset H$, 则称 H 在点 x^0 处正常(非平凡)支撑 S ; 若有 $S \subset H$, 则称 H 在点 x^0 处奇异(平凡)支撑 S .

图 1.3.2 的(a)和(b)分别是正常支撑和奇异支撑的情况.

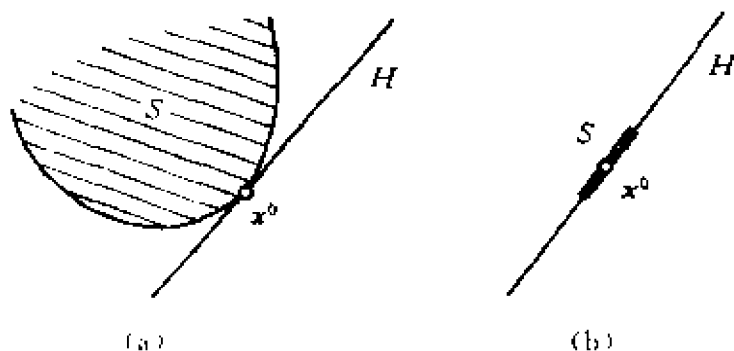


图 1.3.2

推论 1.3.11 设 $S \subset \mathcal{V}$ 是凸集, $S^i \neq \emptyset$, 则对任意的 $x^0 \in S \setminus S^i$, 存在超平面 $H \subset \mathcal{V}$, 它在点 x^0 处正常支撑 S .

证明 设 $S_1 = \{x^0\}$, $S_2 = S$, 由定理的条件有 $S_1^i \cap S_2^i \subset \{x^0\} \cap S^i = \emptyset$. 于是, 利用定理 1.3.5 的(1)得知存在超平面 $H \subset \mathcal{V}$ 分离 $\{x^0\}$ 和 S . 注意到 $\{x^0\}$ 是单点集, 则有 $x^0 \in H$, 又 $S^i \neq \emptyset$, 故 $S \not\subset H$. 因此, 由定义 1.3.3 即得结论. \square

推论 1.3.12 设 $S \subset \mathcal{K}$ 是凸集, $\text{int}S \neq \emptyset$, 则对任意的点 $x^0 \in \text{rb}S$, 存在闭超平面 $H \subset \mathcal{K}$, 它在点 x^0 处正常支撑 S .

证明 与推论 1.3.11 的证明类似, 利用定理 1.3.7 的(1)可得证. \square

定义 1.3.4 设集合 $S \subset \mathcal{V}$ 非空, 并且 $S^i \neq \emptyset$. 记

$$\sigma_S(x^*) = \sup_{x \in S} \langle x^*, x \rangle, \quad x^* \in \mathcal{V}^*,$$

则称 $\sigma_S: \mathcal{V}^* \rightarrow R \cup \{+\infty\}$ 是集合 S 的支撑函数.

定理 1.3.13 设集合 $S \subset \mathcal{X}$ 非空, $\text{ri}S \neq \emptyset$, $\sigma_S: \mathcal{X}^* \rightarrow R \cup \{+\infty\}$ 是集合 S 的支撑函数, 则

$$\text{co}S = \{x \in \mathcal{X} \mid \langle x^*, x \rangle \leq \sigma_S(x^*) \quad \forall x^* \in \mathcal{X}^*\}. \quad (1.3.21)$$

证明 记(1.3.21)右端的集合为 D , 先证 $\text{co}S \subset D$. 不难验证 D 是闭集, 故据定理 1.2.10, 由于 $\text{co}S = \text{cl}(\text{co}S)$, 只需证明 $\text{co}S \subset D$ 即可. 为此, 任取 $x \in \text{co}S$, 由定理 1.1.8 知存在正整数 m , 有 $x = \sum_{i=1}^m \lambda_i x^i$, 其中 $x^i \in S$, $\lambda_i \geq 0$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda_i = 1$. 因 σ_S 是 S 的支撑函数, 由定义 1.3.4 有

$$\langle x^*, x^i \rangle \leq \sup_{x' \in S} \langle x^*, x' \rangle = \sigma_S(x^*), \quad x^* \in \mathcal{X}^*.$$

据此, 得知

$$\begin{aligned} \langle x^*, x \rangle &= \left\langle x^*, \sum_{i=1}^m \lambda_i x^i \right\rangle = \sum_{i=1}^m \lambda_i \langle x^*, x^i \rangle \\ &\leq \sum_{i=1}^m \lambda_i \sigma_S(x^*) = \sigma_S(x^*) \sum_{i=1}^m \lambda_i = \sigma_S(x^*), \end{aligned}$$

从而 $x \in D$.

以下证 $D \subset \text{co}S$. 用反证法, 假设存在 $\bar{x} \in D$, 但 $\bar{x} \notin \text{co}S$, 则从前者有

$$\langle x^*, \bar{x} \rangle \leq \sigma_S(x^*) \quad \forall x^* \in \mathcal{X}^*, \quad (1.3.22)$$

而从后者(由于 $\text{ri}S \neq \emptyset$)得 $\text{ri}(\text{co}S - \{\bar{x}\}) \neq \emptyset$ 和 $\mathbf{0} \notin \text{cl}(\text{co}S - \{\bar{x}\})$. 因此, 由定理 1.3.9 的(1)得知, 存在闭超平面强分离 $\text{co}S$ 和 $\{\bar{x}\}$, 按定义 1.3.2 的(3)也即存在 $\bar{x}^* \in \mathcal{X}^* \setminus \{\mathbf{0}\}$ 使

$$\sup_{x \in \text{co}S} \langle x^*, x \rangle < \langle \bar{x}^*, \bar{x} \rangle.$$

由定义 1.3.4 和上式得

$$\sigma_S(\bar{x}^*) = \sup_{x \in S} \langle \bar{x}^*, x \rangle \leq \sup_{x \in \overline{\text{co}} S} \langle \bar{x}^*, x \rangle < \langle \bar{x}^*, \bar{x} \rangle,$$

这导致与(1.3.22)矛盾. \square

注 1.3.3 若 $S \subset \mathcal{X}$ 是闭凸集, $\text{int} S \neq \emptyset$, 由定理 1.3.13 有

$$S = \{x \in \mathcal{X} \mid \langle x^*, x \rangle \leq \sigma_S(x^*) \quad \forall x^* \in \mathcal{X}^*\}.$$

这表明任一内部为非空的闭凸集均可由它的支撑函数来表示.

对于 $x^* \in \mathcal{X}^*$, 考虑超平面

$$H = \{x \in \mathcal{X} \mid \langle x^*, x \rangle = \sigma_S(x^*)\}. \quad (1.3.23)$$

设有点 $x^0 \in \mathcal{X}$ 使 $\langle x^*, x^0 \rangle = \sigma_S(x^*)$, 则 $x^0 \in H$. 从(1.3.21)和(1.3.23)可知 $S \subset \overline{\text{co}} S \subset H^c$. 因此, 由定义 1.3.3 得知(1.3.23)就是集合 S 在点 x^0 处的支撑超平面.

§ 1.4 凸锥和对偶锥

凸锥和它的对偶锥是两类特殊的凸集. 它们是凸分析和非光滑分析研究中的重要概念和工具.

先介绍线性空间 \mathcal{V} 中的凸锥, 并讨论它的有关性质.

定义 1.4.1 设集合 $K \subset \mathcal{V}$. 若对任意的 $\lambda > 0$, 有

$$\lambda x \in K \quad \forall x \in K,$$

则称 K 是(\mathcal{V} 中的)锥. 若 $\theta \in K$, 则称 K 是有锋锥, 否则, 称 K 是无锋锥.

我们约定: 空集是锥.

定义 1.4.2 设 $K \subset \mathcal{V}$ 是锥.

- (1) 若对任意的 $x \in K \setminus \{\theta\}$, 有 $-x \notin K$, 则称 K 是尖锥.
- (2) 若 K 是闭集, 则称 K 是闭锥.
- (3) 若 K 是凸集, 则称 K 是凸锥.

例 1.4.1 以下集合是相应空间中的凸锥.

Euclid 空间 R^n 中的非负锥:

$$R_+^n = \{x \in R^n \mid x \geq 0\}$$

是 R^n 中的有锋尖闭凸锥; 正锥:

$$\text{int} R_+^n = \{x \in R^n \mid x > 0\}$$

是 R^n 中的无锋尖凸锥.

Banach 空间 \mathcal{B} 中过原点的超平面:

$$H(x^*, 0) = \{x \in \mathcal{B} \mid \langle x^*, x \rangle = 0, x^* \in \mathcal{B}^* \setminus \{0\}\}$$

是 \mathcal{B} 中的有锋闭凸锥, 其中 \mathcal{B}^* 是 \mathcal{B} 的对偶空间.

线性空间 \mathcal{V} 中过原点的开半空间的交:

$$K = \{x \in \mathcal{V} \mid \langle x_i^*, x \rangle > 0, x_i^* \in \mathcal{V}^* \setminus \{0\}, i \in I\}$$

是 \mathcal{V} 中的无锋凸锥 (其中 I 是任意指标集).

定理 1.4.1 设集合 $K \subset \mathcal{V}$. K 是凸锥当且仅当对任意的 $\lambda, \mu > 0$, 有

$$\lambda x^1 + \mu x^2 \in K \quad \forall x^1, x^2 \in K. \quad (1.4.1)$$

证明 必要性. 设 K 是凸锥, 则它是凸集. 于是, 由定义 1.1.1 可知, 对任意的 $\lambda, \mu > 0$, 因为 $\frac{\lambda}{\lambda + \mu}, \frac{\mu}{\lambda + \mu} \in (0, 1)$, 我们有

$$\frac{\lambda}{\lambda + \mu} x^1 + \frac{\mu}{\lambda + \mu} x^2 \in K \quad \forall x^1, x^2 \in K.$$

由于 K 是锥, 故根据定义 1.4.1 得

$$\lambda x^1 + \mu x^2 = (\lambda + \mu) \left(\frac{\lambda}{\lambda + \mu} x^1 + \frac{\mu}{\lambda + \mu} x^2 \right) \in K \quad \forall x^1, x^2 \in K.$$

充分性. 在 (1.4.1) 中取 $x^1 = x^2$, 则对任意的 $\lambda + \mu > 0$, 有 $(\lambda + \mu)x^1 \in K$. 因此, 由定义 1.4.1 知 K 是锥. 取 $\lambda + \mu = 1$, 由 $\lambda,$

$\mu > 0$ 我们有 $\lambda \in (0, 1)$, 并且

$$\lambda x^1 + (1 - \lambda)x^2 \in K \quad \forall x^1, x^2 \in K.$$

于是, 按定义 1.1.1 可知 K 是凸集, 再据定义 1.4.2 的 (3) 使得 K 是凸锥.]

定理 1.4.2 设 $K \subset \mathcal{V}$ 是凸锥, 则 $K + K \subset K$.

证明 设 $x \in K + K$, 则存在 $x^1, x^2 \in K$ 使 $x = x^1 + x^2$. 因为 K 是凸锥, 故由定理 1.4.1 即得 $x = x^1 + x^2 \in K$.]

定理 1.4.3 设 $K_i \subset \mathcal{V}$ ($i = 1, \dots, m$) 是凸锥.

(1) $\bigcap_{i=1}^m K_i$ 是凸锥.

(2) $\sum_{i=1}^m K_i$ 是凸锥.

证明 (1) 任取 $x^1, x^2 \in \bigcap_{i=1}^m K_i$, 则 $x^1, x^2 \in K_i$ ($i = 1, \dots, m$). 因为 K_i ($i = 1, \dots, m$) 是凸锥, 由定理 1.4.1 知, 对任意的 $\lambda, \mu > 0$, 有

$$\lambda x^1 + \mu x^2 \in K_i, \quad i = 1, \dots, m,$$

从而 $\lambda x^1 + \mu x^2 \in \bigcap_{i=1}^m K_i$. 再据定理 1.4.1, 即知 $\bigcap_{i=1}^m K_i$ 是凸锥.

(2) 任取 $x^1, x^2 \in \sum_{i=1}^m K_i$, 则存在 $y_i^1, y_i^2 \in K_i$ ($i = 1, \dots, m$),

使 $x^1 = \sum_{i=1}^m y_i^1, x^2 = \sum_{i=1}^m y_i^2$. 因为 K_i ($i = 1, \dots, m$) 是凸锥, 由定理 1.4.1 知, 对任意的 $\lambda, \mu > 0$, 有

$$\lambda y_i^1 + \mu y_i^2 \in K_i, \quad i = 1, \dots, m.$$

据此, 我们推知

$$\lambda x^1 + \mu x^2 = \sum_{i=1}^m (\lambda y_i^1 + \mu y_i^2) \in \sum_{i=1}^m K_i,$$

再由定理 1.4.1 即得 $\sum_{i=1}^m K_i$ 是凸锥.]

定理 1.4.4 设 \mathscr{V} 是线性空间, $K_1 \subset \mathscr{V}$ 和 $K_2 \subset \mathscr{V}$ 都是凸锥, 则 $K_1 \times K_2 \subset \mathscr{V} \times \mathscr{V}$ 是凸锥.

证明 由定理 1.4.1 可以推证. \square

定理 1.4.5 设 $K_1, K_2 \subset \mathscr{V}$ 是非空凸锥, 则

$$K_1 + K_2 = \text{co}(K_1 \cup K_2).$$

证明 先证 $K_1 + K_2 \subset \text{co}(K_1 \cup K_2)$. 任取 $x \in K_1 + K_2$, 则存在 $x^1 \in K_1, x^2 \in K_2$, 使 $x = x^1 + x^2$. 因为 K_1 和 K_2 是锥, 故有 $2x^1 \in K_1$ 和 $2x^2 \in K_2$, 从而 $2x^1, 2x^2 \in K_1 \cup K_2 \subset \text{co}(K_1 \cup K_2)$. 由于 $\text{co}(K_1 \cup K_2)$ 是凸集, 因此得

$$x = \frac{1}{2}(2x^1) + \frac{1}{2}(2x^2) \in \text{co}(K_1 \cup K_2).$$

现证 $\text{co}(K_1 \cup K_2) \subset K_1 + K_2$. 任取 $x \in \text{co}(K_1 \cup K_2)$, 因为 K_1 和 K_2 是凸集, 由定理 1.1.9 的 (3) 知, 存在 $\bar{\lambda} \in [0, 1]$, 有 $x \in \bar{\lambda}K_1 + (1 - \bar{\lambda})K_2$. 由此, 有 $x^1 \in K_1, x^2 \in K_2$, 使 $x = \bar{\lambda}x^1 + (1 - \bar{\lambda})x^2$. 由于 K_1 和 K_2 是锥, 故 $\bar{\lambda}x^1 \in K_1, (1 - \bar{\lambda})x^2 \in K_2$, 于是得到 $x \in K_1 + K_2$. \square

定义 1.4.3 设集合 $S \subset \mathscr{V}$, 则 \mathscr{V} 中所有包含 S 的有锋凸锥的交称为 S 的锥包 (或由 S 生成的凸锥), 记作 $\text{cone}S$.

我们约定: 空集的锥包是空集.

注 1.4.1 由定义 1.4.3, 集合的锥包是包含它的最小凸锥加上原点. 显然, 若 $K \subset \mathscr{V}$ 是凸锥, 则 $\text{cone}K = K \cup \{0\}$.

定理 1.4.6 设集合 $S \subset \mathscr{V}$, 则对任何正整数 m 有

$$\text{cone}S = \left\{ x \in \mathscr{V} \mid x = \sum_{i=1}^m \lambda_i x^i \quad \forall x^i \in S, \forall \lambda_i \geq 0, i=1, \dots, m \right\},$$

即 $\text{cone}S$ 是 S 中向量的有限非负线性组合构成的集合.

证明 由定义 1.4.3, 利用定理 1.4.1 并注意 $\text{cone}S$ 是包含原点的, 即可推得. \square

定理 1.4.7 设集合 $S_1, S_2 \subset \mathscr{V}$.

(1) $\text{cone}(S_1 \cap S_2) \subset \text{cone}S_1 \cap \text{cone}S_2$. 若 S_1 和 S_2 是凸集并且 $\theta \in S_1 \cap S_2$, 则 $\text{cone}(S_1 \cap S_2) = \text{cone}S_1 \cap \text{cone}S_2$.

(2) $\text{cone}(S_1 + S_2) \subset \text{cone}S_1 + \text{cone}S_2$. 若 $\theta \in S_1 \cap S_2$, 则 $\text{cone}(S_1 + S_2) = \text{cone}S_1 + \text{cone}S_2$.

证明 由定理 1.4.1 可以直接推证. \square

利用凸锥, 我们给出广义的锥凸集和严格锥凸集的概念.

定义 1.4.4 设集合 $S \subset \mathcal{V}$ 非空, $K \subset \mathcal{V}$ 是有锥凸锥. 若对任意的 $\lambda \in (0, 1)$, 有

$$\lambda x^1 + (1 - \lambda)x^2 \in S + K \quad \forall x^1, x^2 \in S,$$

则称 S 是 $(\mathcal{V}$ 中的) K -凸集, 或集合 S 是 K -凸的.

设 \mathcal{X} 是线性拓扑空间.

定义 1.4.5 设集合 $S \subset \mathcal{X}$ 非空, $K \subset \mathcal{V}$ 是内部非空的凸锥. 若对任意的 $\lambda \in (0, 1)$, 有

$$\lambda x^1 + (1 - \lambda)x^2 \in S + \text{int}K \quad \forall x^1, x^2 \in S,$$

则称 S 是 $(\mathcal{X}$ 中的) 严格 K -凸集, 或集合 S 是严格 K -凸的.

注 1.4.2 由定义 1.4.4 和定义 1.4.5 易知, 线性拓扑空间中的严格 K -凸集也是 K -凸集, 但反之不然. 此外, 由定义 1.1.1 和定义 1.4.4 可以得知, 对于某凸锥 K 是 K -凸的集合, 它不一定是凸集, 而凸集关于任何的有锥凸锥 K 都是 K -凸的.

以下介绍线性拓扑空间 \mathcal{X} 中的对偶锥.

设 \mathcal{X}^* 是 \mathcal{X} 的对偶空间, $\langle x^*, x \rangle$ 是连续线性泛函 $x^* \in \mathcal{X}^*$ 在点 $x \in \mathcal{X}$ 处的值.

定义 1.4.6 设 $K \subset \mathcal{X}$ 是凸锥.

(1) 集合

$$K^* = \{x^* \in \mathcal{X}^* \mid \langle x^*, x \rangle \geq 0 \quad \forall x \in K\}$$

称为 K 的对偶锥.

(2) 集合

$$K^{**} = \{x \in \mathcal{X} \mid \langle x, x^* \rangle \geq 0 \quad \forall x^* \in K^*\}$$

称为 K 的双对偶锥.

(3) 集合

$$K^+ = \{x^* \in \mathcal{X}^* \mid \langle x^*, x \rangle > 0 \quad \forall x \in K \setminus \{\theta\}\}$$

称为 K 的严格对偶锥; 集合

$$K^- = \{x^* \in \mathcal{X}^* \mid \langle x^*, x \rangle \leq 0 \quad \forall x \in K\}$$

称为 K 的负对偶锥.

注 1.4.3 由定义 1.4.6 的(1)和(2), 易知 $(K^*)^* = K^{**}$. 因为 $(R^n)^* = R^n$, 特别地有 $(R_+^n)^* = R_+^n$ 和 $(R_+^n)^{**} = R_+^n$. 从定义 1.4.6 中的(1)和(3), 则知

$$K^+ \subset K^*.$$

设 $K_1, K_2 \subset \mathcal{X}$ 是凸锥. 若 $K_1 \subset K_2$, 由定义 1.4.6 的(1)和(3)还不难推知有 $K_2^* \subset K_1^*$ 和 $K_2^- \subset K_1^-$.

定理 1.4.8 设 $K \subset \mathcal{X}$ 是凸锥, 则 K^* 和 K^{**} 是有锋闭凸锥.

证明 先证明 K^* 是有锋闭凸锥. 首先, 由于对任意的 $x \in K$ 有 $\langle \theta, x \rangle = 0$, 按定义 1.4.6 的(1)知 $\theta \in K^*$, 故由定义 1.4.1 得 K^* 是有锋锥. 其次, 设序列 $\{x_k^*\} \subset K^*$, $x_k^* \rightarrow x^* (k \rightarrow \infty)$, 则由定义 1.4.6 的(1), 对任意的 $x \in K$ 有

$$\langle x_k^*, x \rangle \geq 0, \quad k = 1, 2, \dots.$$

对上式令 $k \rightarrow \infty$, 因为 $x_k^* \in \mathcal{X}^* (k = 1, 2, \dots)$ 是 x 处的连续泛函, 所以由 $x_k^* \rightarrow x^*$ 得 $\langle x^*, x \rangle \geq 0$. 由此, 根据定义 1.4.6 的(1)知 $x^* \in K^*$, 于是 K^* 是闭集. 最后, 为证 K^* 是凸锥, 任取 $x^1, x^2 \in K^*$, 由定义 1.4.6 的(1), 对任意的 $x \in K$ 有 $\langle x^1, x \rangle \geq 0$ 和 $\langle x^2, x \rangle \geq 0$. 由于 $x^1, x^2 \in \mathcal{X}^*$ 又是 x 处的线性泛函, 故对任意的 $\lambda, \mu > 0$ 有

$$\langle \lambda x^1 + \mu x^2, x \rangle = \lambda \langle x^1, x \rangle + \mu \langle x^2, x \rangle \geqslant 0.$$

据此,由定义 1.4.6 的(1)知 $\lambda x^1 + \mu x^2 \in K^*$,再由定理 1.4.1 即得 K^* 是凸锥.

与上面的证明类似,由定义 1.4.6 的(2)注意到 $x^* \in \mathcal{X}^*$ 是连续线性泛函,可得 K^{**} 是有锋闭凸锥. \square

定理 1.4.9 设 \mathcal{X} 是分离局部凸线性拓扑空间, $K \subset \mathcal{X}$ 是凸锥.

(1) $K^{**} = \text{cl}K$.

(2) 若 K 是闭集,则 $K^{**} = K$.

证明 (1) 任取 $x \in K$, 由定义 1.4.6 的(1),对任意的 $x^* \in K^*$ 有 $\langle x^*, x \rangle \geqslant 0$. 于是,再由定义 1.4.6 的(2)知 $x \in K^{**}$,从而 $K \subset K^{**}$. 根据定理 1.4.8, K^{**} 是闭集,因而 $\text{cl}K \subset K^{**}$. 为证 $K^{**} \subset \text{cl}K$, 设 $x \notin \text{cl}K$, 则因 $\text{cl}K$ 是闭凸集,故由定理 1.3.9 知点 x 和集合 $\text{cl}K$ 可强分离. 于是,按定义 1.3.2 的(3),存在 $x^* \in \mathcal{X}^*$ 使得

$$\langle x^*, x \rangle < \inf_{y \in \text{cl}K} \langle x^*, y \rangle. \quad (1.4.2)$$

显然有 $\theta \in \text{cl}K$, 则 $\inf_{y \in \text{cl}K} \langle x^*, y \rangle \leqslant 0$, 因而由(1.4.2)得知

$$\langle x^*, x \rangle < 0. \quad (1.4.3)$$

设若存在 $\bar{y} \in K \setminus \{\theta\}$ 使 $\langle x^*, \bar{y} \rangle < 0$, 则有 $\langle x^*, \lambda \bar{y} \rangle \rightarrow -\infty$ ($\lambda \rightarrow +\infty$), 从而得 $\inf_{y \in \text{cl}K} \langle x^*, y \rangle = -\infty$, 与(1.4.2)相矛盾. 由此,对任意的 $y \in K \setminus \{\theta\}$ 有 $\langle x^*, y \rangle \geqslant 0$, 于是

$$\langle x^*, y \rangle \geqslant 0 \quad \forall y \in K.$$

按定义 1.4.6 的(1)知 $x^* \in K^*$, 再由(1.4.3)便得 $x \notin K^{**}$.

(2) 由 K 是闭集,按定义 1.2.2 的(3)有 $K = \text{cl}K$. 于是,从已证得的(1)即得结论. \square

注 1.4.4 设 K 是分离局部凸空间 \mathcal{X} 中的凸锥,与定理 1.4.9 的证明类似,可以得到 $(K^-)^- = \text{cl}K$. 再若 K 是闭的,则

$$(K^{\circ})^{\circ} = K.$$

定理 1.4.10 设 $K_1, K_2 \subset \mathcal{X}$ 是凸锥.

(1) 若 $K_1 \subset K_2$, 则 $K_2^* \subset K_1^*$.

(2) $(K_1 + K_2)^* = K_1^* \cap K_2^* = (K_1 \cup K_2)^*$.

(3) $\text{co}(K_1^* \cup K_2^*) = K_1^* + K_2^*$.

(4) $(\text{cl}K_1 \cap \text{cl}K_2)^* = \text{cl}(K_1^* + K_2^*)$.

证明 (1) 由定义 1.4.6 的(1)直接可得.

(2) 先证 $(K_1 + K_2)^* = K_1^* \cap K_2^*$. 设 $x^* \in (K_1 + K_2)^*$, 由定义 1.4.6 的(1)有

$$\langle x^*, x^1 + x^2 \rangle \geq 0 \quad \forall x^1 \in K_1, \forall x^2 \in K_2.$$

在上式中分别令 x^1 和 x^2 趋于 θ , 则有

$$\langle x^*, x^1 \rangle \geq 0 \quad \forall x^1 \in K_1$$

和

$$\langle x^*, x^2 \rangle \geq 0 \quad \forall x^2 \in K_2.$$

由此, 得到 $x^* \in K_1^*$ 和 $x^* \in K_2^*$, 从而 $x^* \in K_1^* \cap K_2^*$. 因上述过程可逆, 故得证.

下证 $K_1^* \cap K_2^* = (K_1 \cup K_2)^*$. 设 $x^* \in (K_1 \cup K_2)^*$, 由定义 1.4.6 的(1)有

$$\langle x^*, x \rangle \geq 0 \quad \forall x \in K_1 \cup K_2.$$

于是得知

$$\langle x^*, x \rangle \geq 0 \quad \forall x \in K_1$$

和

$$\langle x^*, x \rangle \geq 0 \quad \forall x \in K_2,$$

从而 $x^* \in K_1^* \cap K_2^*$. 上述过程也可逆, 故得证.

(3) 因为 K_1 和 K_2 是凸锥, 由定理 1.4.8 知, K_1^* 和 K_2^* 是线性空间 \mathcal{X}^* 中的非空凸锥. 由此, 根据定理 1.4.5 即得结论.

(4) 由定理 1.4.9 的(1), 有

$$K_1^{**} = \text{cl}K_1, \quad K_2^{**} = \text{cl}K_2.$$

再利用已证得的(2)和定理 1.4.9 的(1), 可得

$$\begin{aligned} (\text{cl}K_1 \cap \text{cl}K_2)^* &= (K_1^{**} \cap K_2^{**})^* = [(K_1^* + K_2^*)^*]^* \\ &= (K_1^* + K_2^*)^{**} = \text{cl}(K_1^* + K_2^*). \quad \square \end{aligned}$$

定理 1.4.11 设 $K \subset \mathcal{K}$ 是凸锥, $\text{int}K \neq \emptyset$.

(1) 若 $x^* \in K^* \setminus \{\theta\}$, $x \in \text{int}K$, 则 $\langle x^*, x \rangle > 0$.

(2) 若 $x^* \in \text{int}K^*$, $x \in K \setminus \{\theta\}$, 则 $\langle x^*, x \rangle > 0$.

证明 (1) 用反证法. 假设存在 $\bar{x}^* \in K^* \setminus \{\theta\}$ 和 $\bar{x} \in \text{int}K$, 有

$$\langle \bar{x}^*, \bar{x} \rangle \leq 0. \quad (1.4.4)$$

由 $\text{int}K$ 是开集和 $\bar{x} \in \text{int}K$, 可知存在 \bar{x} 的邻域 $U(\bar{x}) \subset K$, 由定义 1.4.6 的(1)有

$$\langle x^*, x \rangle \geq 0 \quad \forall x^* \in K^*, \forall x \in U(\bar{x}).$$

因为 $\bar{x}^* \in K^* \setminus \{\theta\} \subset K^*$, 所以从上式得

$$\langle \bar{x}^*, x \rangle \geq 0 \quad \forall x \in U(\bar{x}). \quad (1.4.5)$$

又由 $\bar{x}^* \neq \theta$ 知存在 $x^0 \in \mathcal{K}$, 有 $\langle \bar{x}^*, x^0 \rangle > 0$. 取充分小的 $\varepsilon > 0$, 使 $x = \bar{x} - \varepsilon x^0 \in U(\bar{x})$, 则由(1.4.5)得

$$\langle \bar{x}^*, \bar{x} - \varepsilon x^0 \rangle \geq 0.$$

于是推知 $\langle \bar{x}^*, \bar{x} \rangle \geq \varepsilon \langle \bar{x}^*, x^0 \rangle > 0$, 这与(1.4.4)相矛盾.

(2) 与(1)的证明类似, 利用 $\text{int}K^*$ 是开集即可证明. \square

第2章 凸函数和凸映射

凸函数和凸映射是凸分析的主要研究对象,其有关理论也是非光滑分析的重要基础.

本章先阐述凸函数的概念及其基本性质,然后讨论它的连续性和对偶性问题.在论述了两类广义意义下的凸函数之后,进而将函数的凸性概念推广到映射凸性的情况,介绍凸映射和广义凸映射.

§ 2.1 凸函数及有关性质

我们对取实数值的函数和实泛函统称为实值函数.这一节介绍线性空间上具有凸性的实值函数以及它的有关性质.

设 \mathcal{V} 是线性空间.

定义 2.1.1 设 $S \subset \mathcal{V}$ 是非空凸集, $f: S \rightarrow R$ 是实值函数.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f(x^1) + (1 - \lambda)f(x^2) \\ \forall x^1, x^2 \in S, \quad (2.1.1)$$

则称 f 是集合 S 上的凸函数(凸泛函),或函数(泛函) f 在 S 上是凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda x^1 + (1 - \lambda)x^2) < \lambda f(x^1) + (1 - \lambda)f(x^2) \\ \forall x^1, x^2 \in S, f(x^1) \neq f(x^2), \quad (2.1.2)$$

则称 f 是集合 S 上的严格凸函数, 或函数 f 在 S 上是严格凸的.

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda x^1 + (1 - \lambda)x^2) < \lambda f(x^1) + (1 - \lambda)f(x^2)$$

$$\forall x^1, x^2 \in S, x^1 \neq x^2, \quad (2.1.3)$$

则称 f 是集合 S 上的强凸函数, 或函数 f 在 S 上是强凸的.

若 $-f$ 是 S 上的凸函数, 则称 f 是 S 上的凹函数, 或 f 在 S 上是凹的. 若 $-f$ 是 S 上的严格(强)凸函数, 则称 f 是 S 上的严格(强)凹函数, 或 f 在 S 上是严格(强)凹的.

图 2.1.1 中的(a)、(b)、(c)和(d)依次是单变量的凸函数、严格凸函数、强凸函数和凹函数的图示.

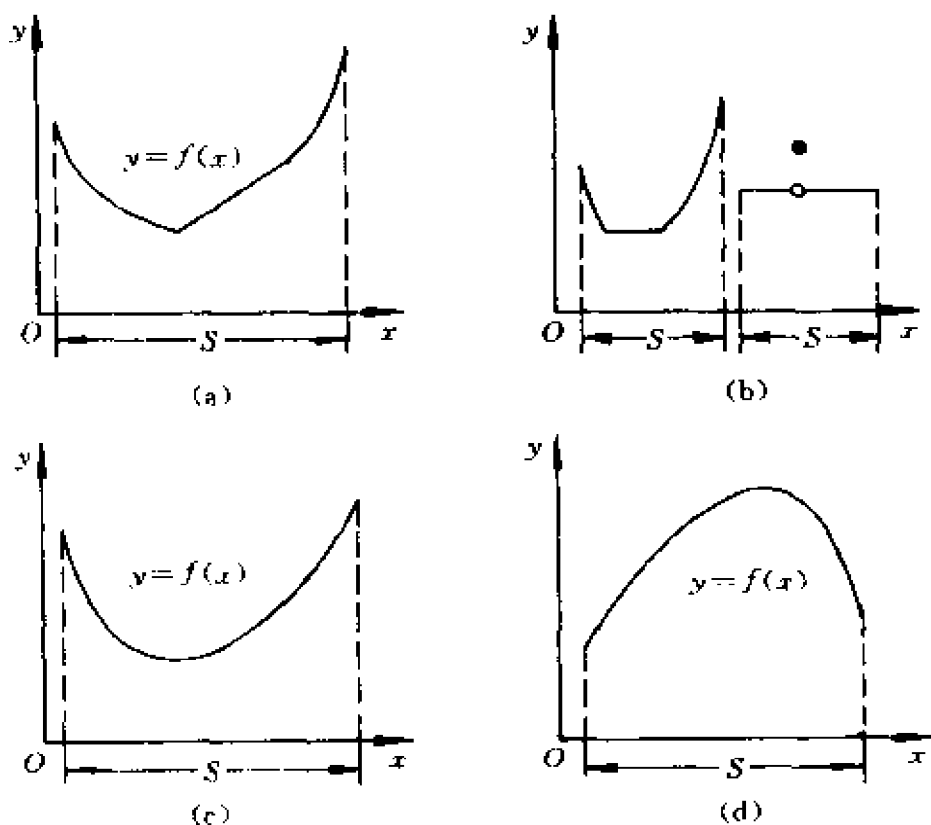


图 2.1.1

注 2.1.1 在定义 2.1.1 中固定 $x^1 = x^0 \in S$, 令 $x^2 = x$, 则 (2.1.1) 成为

$$f(\lambda x^0 + (1 - \lambda)x) \leq \lambda f(x^0) + (1 - \lambda)f(x) \quad \forall x \in S.$$

这时, 称 f 在点 x^0 处相对于集合 S 是凸的. 同样地, (2.1.2) (或 (2.1.3)) 成为

$$f(\lambda x^0 + (1 - \lambda)x) < \lambda f(x^0) + (1 - \lambda)f(x)$$

$$\forall x \in S, f(x^0) \neq f(x) \text{ (或 } x^0 \neq x),$$

并称 f 在点 x^0 处相对于 S 是严格凸的 (或强凸的).

注 2.1.2 从定义 2.1.1 易知, 若 f 是 S 上的强凸 (凹) 函数, 则 f 必是 S 上的凸 (凹) 函数和严格凸 (凹) 函数, 但反之不然. 注意, 严格凸函数不一定是凸函数 (见图 2.1.1(b)).

例 2.1.1 以下函数是相应集合上的凸函数.

实数域 R 上的二次函数:

$$f(x) = x^2, \quad x \in R,$$

并且是严格凸函数和强凸函数.

Euclid 空间 R^n 上的范数函数:

$$f(x) = \|x\|_p = \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}}, \quad x \in R^n, p > 1,$$

其中 $x = (x_1, \dots, x_n)^T$, 特别地,

$$f(x) = \|x\| = \sqrt{x_1^2 + \dots + x_n^2}$$

是 R^n 上的凸函数.

Banach 空间 \mathcal{B} 中凸集 S 上的距离函数:

$$d_S(x) = \inf_{y \in S} \|x - y\|, \quad x \in \mathcal{B}.$$

线性拓扑空间 \mathcal{X} 中凸集 S 上的 Minkowski 函数 (泛函):

$$\mu_S(x) = \inf \{ \alpha > 0 \mid x \in \alpha S \}, \quad x \in \mathcal{X}.$$

线性空间 \mathcal{V} 上的仿射函数:

$$l(x) = \langle a, x \rangle + \beta, \quad x \in \mathcal{V},$$

其中 $a \in \mathcal{V}$, $\beta \in R$. 它也是 \mathcal{V} 上的凹函数, 而不是 \mathcal{V} 上的严格凸函数.

线性空间 \mathcal{V} 中凸集 S 上的指示函数:

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathcal{V} \setminus S. \end{cases}$$

先给出关于凸函数的两个基本关系.

定理 2.1.1 设 $S \subset \mathcal{V}$ 是非空凸集, 则 $f: S \rightarrow R$ 是凸函数当且仅当对任何正整数 $m \geq 2$ 和任意的 $\lambda_i \geq 0$ ($i = 1, \dots, m$),

$\sum_{i=1}^m \lambda_i = 1$, 有

$$f\left(\sum_{i=1}^m \lambda_i x^i\right) \leq \sum_{i=1}^m \lambda_i f(x^i) \quad \forall x^1, \dots, x^m \in S. \quad (2.1.4)$$

证明 充分性. 在 (2.1.4) 中取 $m = 2$, 由定义 2.1.1 中的 (1) 即得.

必要性. 用数学归纳法证. 当 $m = 2$ 时, 因为 f 是 S 上的凸函数, 由 (2.1.1) 知 (2.1.4) 成立. 现设 $m \leq k$ ($k \geq 2$ 是正整数) 时 (2.1.4) 成立, 证明 (2.1.4) 对于 $m = k + 1$ 成立. 不失一般性, 设 $\lambda_i > 0$ ($i = 1, \dots, k + 1$), $\sum_{i=1}^{k+1} \lambda_i = 1$ (若有某 $\lambda_i = 0$, 则 $m = k$,

由假设知 (2.1.4) 成立), 则有 $1 - \lambda_{k+1} = \sum_{i=1}^k \lambda_i > 0$. 记 $\sum_{i=1}^k \lambda_i x^i =$

$(1 - \lambda_{k+1})x$, 由 $x^i \in S$ ($i = 1, \dots, k$) 和 $\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} = 1$ 以及 S

是凸集, 根据定理 1.1.4 有 $x = \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \in S$. 又因为 $x^{k+1} \in S$, 以及 f 是 S 上的凸函数, 按定义 2.1.1 的 (1), 得知

$$f\left(\sum_{i=1}^{k+1} \lambda_i x^i\right) = f\left((1 - \lambda_{k+1})x + \lambda_{k+1}x^{k+1}\right)$$

$$\begin{aligned} &\leqslant (1 - \lambda_{k+1})f(x) + \lambda_{k+1}f(x^{k+1}) \\ &= (1 - \lambda_{k+1})f\left(\frac{1}{1 - \lambda_{k+1}}\sum_{i=1}^k \lambda_i x^i\right) + \lambda_{k+1}f(x^{k+1}). \end{aligned}$$

由于假设(2.1.4)对于 $m = k$ 成立,故有

$$f\left(\frac{1}{1 - \lambda_{k+1}}\sum_{i=1}^k \lambda_i x^i\right) \leqslant \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} f(x^i).$$

将它代入上式,得到

$$\begin{aligned} f\left(\sum_{i=1}^{k+1} \lambda_i x^i\right) &\leqslant (1 - \lambda_{k+1}) \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} f(x^i) + \lambda_{k+1}f(x^{k+1}) \\ &= \sum_{i=1}^{k+1} \lambda_i f(x^i), \end{aligned}$$

即(2.1.4)对于 $m = k + 1$ 成立. \square

定理 2.1.2 设 $S \subset \mathscr{X}$ 是非空凸集, $f: S \rightarrow R$ 是凸函数,则对任意的 $c \in R$, 水平集 $H_S(f, c) = \{x \in S \mid f(x) \leqslant c\}$ 是凸集.

证明 任取 $x^1, x^2 \in H_S(f, c) \subset S$, 因为 S 是凸集,故对任意的 $\lambda \in (0, 1)$ 有 $\lambda x^1 + (1 - \lambda)x^2 \in S$. 由于 f 是 S 上的凸函数,按定义 2.1.1 的(1),得

$$\begin{aligned} f(\lambda x^1 + (1 - \lambda)x^2) &\leqslant \lambda f(x^1) + (1 - \lambda)f(x^2) \\ &\leqslant \lambda c + (1 - \lambda)c = c, \end{aligned}$$

于是 $\lambda x^1 + (1 - \lambda)x^2 \in H_S(f, c)$. 据此,由定义 1.1.1 即知 $H_S(f, c)$ 是凸集. \square

为了讨论凸函数和严格凸函数之间的关系,我们引入实值函数(实泛函)的半连续性概念如下.

定义 2.1.2 设 \mathscr{X} 是线性拓扑空间, $f: \mathscr{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数,点 $x^0 \in \mathscr{X}$.

(1) 若对任意的 $\varepsilon > 0$, 存在 x^0 的邻域 $U(x^0)$ 有

$$- \varepsilon < f(x) - f(x^0) \quad \forall x \in U(x^0),$$

则称 f 在点 x^0 处是下半连续的, 若 f 在集合 $S \subset \mathscr{X}$ 的每一点处都是下半连续的, 则称 f 是 S 上的下半连续函数(泛函), 或 f 在 S 上是下半连续的.

(2) 若对任意的 $\varepsilon > 0$, 存在 x^0 的邻域 $U(x^0)$ 有

$$f(x) - f(x^0) < \varepsilon \quad \forall x \in U(x^0),$$

则称 f 在点 x^0 处是上半连续的, 若 f 在集合 $S \subset \mathscr{X}$ 的每一点处都是上半连续的, 则称 f 是 S 上的上半连续函数(泛函), 或 f 在 S 上是上半连续的.

注 2.1.3 由定义 2.1.2 的(1), f 在点 x^0 处是下半连续的等价于

$$f(x^0) - \varepsilon \leq \inf_{x \in U(x^0)} f(x) \leq \sup_{U(x^0)} \inf_{x \in U(x^0)} f(x),$$

令 $\varepsilon \rightarrow 0$ 即

$$f(x^0) \leq \liminf_{x \rightarrow x^0} f(x),$$

其中 $\liminf_{x \rightarrow x^0} f(x) = \sup_{U(x^0)} \inf_{x \in U(x^0)} f(x)$ 是 f 在点 x^0 处的下极限^[14], 同理, 由定义 2.1.2 的(2), f 在点 x^0 处是上半连续的等价于

$$f(x^0) \geq \limsup_{x \rightarrow x^0} f(x),$$

其中 $\limsup_{x \rightarrow x^0} f(x) = \inf_{U(x^0)} \sup_{x \in U(x^0)} f(x)$ 是 f 在点 x^0 处的上极限.

定理 2.1.3 设 \mathscr{X} 是线性拓扑空间, $S \subset \mathscr{X}$ 是非空凸集. 若 $f: S \rightarrow R$ 是下半连续的严格凸函数, 则 f 是 S 上的凸函数.

证明 设对任意的 $x^1, x^2 \in S$, 有 $f(x^1) \neq f(x^2)$. 由 f 在 S 上是严格凸的, 从(2.1.2)知(2.1.1)成立. 因此, f 在 S 上是凸的.

现设存在 $\bar{x}^1, \bar{x}^2 \in S$, 有 $f(\bar{x}^1) = f(\bar{x}^2)$. 用反证法, 假设 f 在 S 上不是凸的, 则由定义 2.1.1 的(1)知, 存在 $\lambda_0 \in (0, 1)$, 有

$$f(\lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2) > \lambda_0 f(\bar{x}^1) + (1 - \lambda_0) f(\bar{x}^2)$$

$$= f(\bar{x}^1) = f(\bar{x}^2).$$

记 $x^0 = \lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2$, 由 S 是凸集知 $x^0 \in S$, 并从上式得到

$$f(\bar{x}^1) < f(x^0), \quad f(\bar{x}^2) < f(x^0). \quad (2.1.5)$$

对于(2.1.5)的第1式, 取充分小的 $\epsilon_0 > 0$, 可使 $f(\bar{x}^1) + \epsilon_0 < f(x^0)$, 或

$$f(\bar{x}^1) < f(x^0) - \epsilon_0.$$

因为 f 在 $x^0 \in S$ 处是下半连续的, 由定义 2.1.2 的(1)知, 对任意的 $\epsilon > 0$, 存在 λ_0 的邻域 $U(\lambda_0)$ 有

$$-\epsilon < f[\lambda \bar{x}^1 + (1 - \lambda) \bar{x}^2] - f(x^0) \quad \forall \lambda \in U(\lambda_0),$$

从而得

$$f(\bar{x}^1) < f(x^0) - \epsilon_0 < f[\lambda \bar{x}^1 + (1 - \lambda) \bar{x}^2] \quad \forall \lambda \in U(\lambda_0).$$

取 $\lambda' \in U(\lambda_0)$, $\lambda' \neq \lambda_0$, 记 $x' = \lambda' \bar{x}^1 + (1 - \lambda') \bar{x}^2$. 从上式有

$$f(\bar{x}^1) < f(x'). \quad (2.1.6)$$

不妨设 $x^0 \in (\bar{x}^1, x')$, 则 $x' \in (x^0, \bar{x}^2)$ (若设 $x^0 \in (x', \bar{x}^2)$, 则 $x' \in (\bar{x}^1, x^0)$, 同理可推证), 于是存在 $\lambda_1, \lambda_2 \in (0, 1)$, 有 $x^0 = \lambda_1 \bar{x}^1 + (1 - \lambda_1)x'$ 和 $x' = \lambda_2 x^0 + (1 - \lambda_2) \bar{x}^2$. 由于 f 在 S 上是严格凸的, 依据定义 2.1.1 的(2)并注意到(2.1.6)和(2.1.5)的第2式, 可得

$$\begin{aligned} f(x^0) &= f(\lambda_1 \bar{x}^1 + (1 - \lambda_1)x') \\ &< \lambda_1 f(\bar{x}^1) + (1 - \lambda_1)f(x') < f(x') \end{aligned}$$

和

$$\begin{aligned} f(x') &= f(\lambda_2 x^0 + (1 - \lambda_2) \bar{x}^2) \\ &< \lambda_2 f(x^0) + (1 - \lambda_2)f(\bar{x}^2) < f(x^0). \end{aligned}$$

导致以上两式相矛盾. \square

为了把凸函数的值域扩展到广义实值域 $R \cup \{\pm \infty\}$, 我们对涉及 $\pm \infty$ 的运算作如下约定:

$$\left\{ \begin{array}{l} \alpha + \infty = +\infty + \alpha = +\infty \quad \forall \alpha \in R \cup \{+\infty\}, \\ \alpha + (-\infty) = -\infty + \alpha = -\infty \quad \forall \alpha \in R \cup \{-\infty\}, \\ +\infty + (+\infty) = +\infty, \quad -\infty + (-\infty) = -\infty, \\ \lambda \cdot (+\infty) = (+\infty) \cdot \lambda = +\infty, \\ \quad \lambda \cdot (-\infty) = (-\infty) \cdot \lambda = -\infty \quad \forall \lambda > 0, \\ \lambda \cdot (+\infty) = (+\infty) \cdot \lambda = -\infty, \\ \quad \lambda \cdot (-\infty) = (-\infty) \cdot \lambda = +\infty \quad \forall \lambda < 0, \\ 0 \cdot (+\infty) = (+\infty) \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0 = 0, \\ +\infty + (+\infty) = +\infty, \\ -(-\infty) = +\infty, \\ \inf \emptyset = +\infty, \sup \emptyset = -\infty. \end{array} \right. \quad (2.1.7)$$

对于(2.1.7), 还规定: 只允许使用“消去”而不可使用“移项”. 有了以上规定, 我们就可以把任何凸集上的凸函数扩展到全空间.

定义 2.1.3 设 $S \subset \mathcal{V}$ 是非空凸集, $\hat{f}: S \rightarrow R$ 是凸函数. 令

$$f(x) = \begin{cases} \hat{f}(x), & x \in S; \\ +\infty, & x \in \mathcal{V} \setminus S, \end{cases}$$

则称 $f: \mathcal{V} \rightarrow R \cup \{\pm \infty\}$ 是空间 \mathcal{V} 上的广义实值凸函数, 也简称凸函数.

依照涉及 $+\infty$ 和 $-\infty$ 的约定运算, 定义 2.1.3 和定义 2.1.1 的(1)是一致的.

定义 2.1.4 设 $f: \mathcal{V} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 则称集合

$$\text{dom} f = \{x \in \mathcal{V} \mid f(x) < +\infty\}$$

是 f 的有效域. 若 $\text{dom} f \neq \emptyset$, 则称 f 是正常函数, 或函数 f 是正常的.

定义 2.1.5 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数. 若 $\text{dom} f \neq \emptyset$, 并且对于任意的 $x \in \text{dom} f$, 有 $f(x) \neq -\infty$, 则称 f 是 ($\text{dom} f$ 上的) 正常凸函数, 或函数 f 是正常凸的. 否则, 称 f 是非正常凸函数. 若 $-f$ 是正常凸函数, 则称 f 是正常凹函数.

注 2.1.4 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数, 由定义 2.1.3、定义 2.1.1 的(1)和定义 2.1.4, 不难得知 $\text{dom} f$ 是凸集.

定理 2.1.4 设 $f_i: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ ($i = 1, 2$) 是正常凸函数.

$$(1) \text{dom}(f_1 + f_2) = \text{dom} f_1 \cap \text{dom} f_2.$$

(2) 若 $\text{dom} f_1 \cap \text{dom} f_2 \neq \emptyset$, 则 $f_1 + f_2$ 是正常凸函数.

证明 (1) 设 $x \in \text{dom}(f_1 + f_2)$, 由定义 2.1.4 有 $f_1(x) + f_2(x) < +\infty$. 又因为 f_1 和 f_2 是正常凸函数, 由定义 2.1.5 知 $f_1(x)$ 和 $f_2(x)$ 均不取 $-\infty$. 由此推知, 有 $f_1(x) < +\infty$ 和 $f_2(x) < +\infty$, 于是得 $x \in \text{dom} f_1 \cap \text{dom} f_2$. 反之亦然.

(2) 记 $f(x) = f_1(x) + f_2(x)$. 任取 $x^1, x^2 \in \mathcal{X}$ 和任意的 $\lambda \in (0, 1)$, 因为 f_1 和 f_2 是凸函数, 从而有

$$f_i(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f_i(x^1) + (1 - \lambda)f_i(x^2) \quad (i = 1, 2).$$

将上式对 i 相加, 得

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f(x^1) + (1 - \lambda)f(x^2),$$

因此 f 是凸函数. 从(1)知

$$\text{dom} f = \text{dom}(f_1 + f_2) = \text{dom} f_1 \cap \text{dom} f_2 \neq \emptyset,$$

故对任意的 $x \in \text{dom} f$, 有 $x \in \text{dom} f_1$ 和 $x \in \text{dom} f_2$. 由 f_1 和 f_2 是正常凸函数, 得 $f_1(x) \neq -\infty$ 和 $f_2(x) \neq -\infty$. 于是, $f = f_1 + f_2$ 是正常凸函数. \square

现在引进广义实值函数的上图象和下图象, 并讨论当它们是凸集时分别与凸函数和凹函数的关系.

定义 2.1.6 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数.

(1) 集合

$$\text{epi} f = \{(x, \eta) \in \mathcal{V} \times R \mid \eta \geq f(x)\}$$

称为 f 的上图象.

(2) 集合

$$\text{hyp} f = \{(x, \eta) \in \mathcal{V} \times R \mid \eta \leq f(x)\}$$

称为 f 的下图象.

图 2.1.2 的(a)和(b)分别是单变量广义实值函数的上图象和下图象的图示.

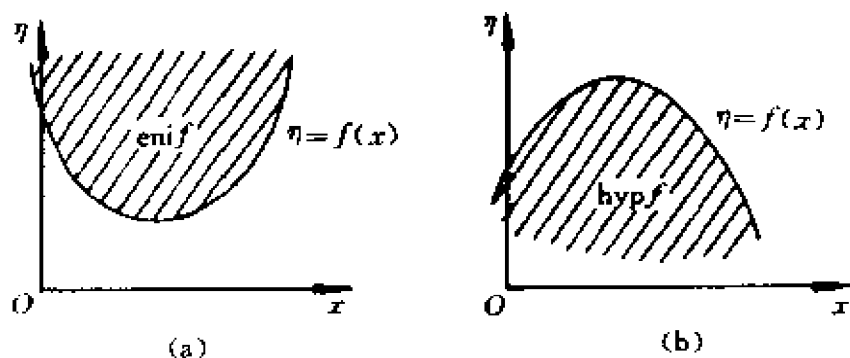


图 2.1.2

定理 2.1.5 设 $f: \mathcal{V} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数. 以下条件是等价的:

- (1) f 是凸函数;
- (2) $\text{epi} f$ 是凸集;
- (3) $(\text{epi} f)^+ = \{(x, \eta) \in \mathcal{V} \times R \mid \eta > f(x)\}$ 是凸集.

证明 (1) \Rightarrow (2). 任取 $(x^1, \eta_1), (x^2, \eta_2) \in \text{epi} f$, 由定义 2.1.6 的(1)有

$$\eta_1 \geq f(x^1), \eta_2 \geq f(x^2).$$

对于任意的 $\lambda \in (0, 1)$, 因为 f 是凸的, 则得

$$\begin{aligned} f(\lambda x^1 + (1 - \lambda)x^2) &\leq \lambda f(x^1) + (1 - \lambda)f(x^2) \\ &\leq \lambda \eta_1 + (1 - \lambda)\eta_2. \end{aligned}$$

据此,再由定义 2.1.6 的(1)有

$$\begin{aligned} & \lambda(\mathbf{x}^1, \eta_1) + (1 - \lambda)(\mathbf{x}^2, \eta_2) \\ &= (\lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2, \lambda\eta_1 + (1 - \lambda)\eta_2) \in \text{epi}f, \end{aligned}$$

于是按定义 1.1.1 知 $\text{epi}f$ 是凸集.

(2) \Rightarrow (3). 任取 $(\mathbf{x}^1, \eta_1), (\mathbf{x}^2, \eta_2) \in (\text{epi}f)^+$, 则

$$\eta_1 > f(\mathbf{x}^1), \eta_2 > f(\mathbf{x}^2).$$

由此可知,存在 $\xi_1, \xi_2 \in R$, 有 $\eta_1 > \xi_1 \geq f(\mathbf{x}^1), \eta_2 > \xi_2 \geq f(\mathbf{x}^2)$, 从而 $(\mathbf{x}^1, \xi_1), (\mathbf{x}^2, \xi_2) \in \text{epi}f$. 因为 $\text{epi}f$ 是凸集,由定义 1.1.1 知,对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} & (\lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2, \lambda\xi_1 + (1 - \lambda)\xi_2) \\ &= \lambda(\mathbf{x}^1, \xi_1) + (1 - \lambda)(\mathbf{x}^2, \xi_2) \in \text{epi}f, \end{aligned}$$

于是由定义 2.1.6 的(1)得

$$\lambda\eta_1 + (1 - \lambda)\eta_2 > \lambda\xi_1 + (1 - \lambda)\xi_2 \geq f(\lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2).$$

这说明

$$\begin{aligned} & \lambda(\mathbf{x}^1, \eta_1) + (1 - \lambda)(\mathbf{x}^2, \eta_2) \\ &= (\lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2, \lambda\eta_1 + (1 - \lambda)\eta_2) \in (\text{epi}f)^+, \end{aligned}$$

于是推得 $(\text{epi}f)^+$ 是凸集.

(3) \Rightarrow (1). 任取 $\mathbf{x}^1, \mathbf{x}^2 \in \mathcal{X}$, 则对任意的 $\varepsilon > 0$ 有 $f(\mathbf{x}^1) + \varepsilon > f(\mathbf{x}^1), f(\mathbf{x}^2) + \varepsilon > f(\mathbf{x}^2)$, 即有 $(\mathbf{x}^1, f(\mathbf{x}^1) + \varepsilon), (\mathbf{x}^2, f(\mathbf{x}^2) + \varepsilon) \in (\text{epi}f)^+$. 由于 $(\text{epi}f)^+$ 是凸集,故对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} & (\lambda\mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2, \lambda(f(\mathbf{x}^1) + \varepsilon) + (1 - \lambda)(f(\mathbf{x}^2) + \varepsilon)) \\ &= \lambda(\mathbf{x}^1, f(\mathbf{x}^1) + \varepsilon) + (1 - \lambda)(\mathbf{x}^2, f(\mathbf{x}^2) + \varepsilon) \in (\text{epi}f)^+, \end{aligned}$$

于是有

$$\lambda f(\mathbf{x}^1) + (1 - \lambda)f(\mathbf{x}^2) + \varepsilon > f(\lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2).$$

令 $\varepsilon \rightarrow 0$, 得到

$$f(\lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2) \leqslant \lambda f(\mathbf{x}^1) + (1 - \lambda)f(\mathbf{x}^2),$$

因而 f 是凸函数. \square

注 2.1.5 与定理 2.1.5 类似, f 是凹函数与 $\text{hyp}f$ 是凸集等价.

下面介绍凸函数的有关性质.

定理 2.1.6 设 $f_i: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数, $\alpha_i \geqslant 0$ ($i = 1, \dots, m$). 若

$$f(\mathbf{x}) = \sum_{i=1}^m \alpha_i f_i(\mathbf{x}),$$

则 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数.

证明 取任意的 $\mathbf{x}^1, \mathbf{x}^2 \in \mathcal{X}$ 和任意的 $\lambda \in (0, 1)$, 因为 f_i ($i = 1, \dots, m$) 是凸函数, 由定义 2.1.3 和定义 2.1.1 的(1), 有

$$\begin{aligned} f(\lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2) &= \sum_{i=1}^m \alpha_i f_i(\lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2) \\ &\leqslant \sum_{i=1}^m \alpha_i [\lambda f_i(\mathbf{x}^1) + (1 - \lambda)f_i(\mathbf{x}^2)] \\ &= \lambda \sum_{i=1}^m \alpha_i f_i(\mathbf{x}^1) + (1 - \lambda) \sum_{i=1}^m \alpha_i f_i(\mathbf{x}^2) \\ &= \lambda f(\mathbf{x}^1) + (1 - \lambda)f(\mathbf{x}^2). \end{aligned}$$

据此, 即知 f 是凸函数. \square

定理 2.1.7 设 $f_i: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数 ($i \in I$, I 是指标集). 若

$$f(\mathbf{x}) = \sup_{i \in I} f_i(\mathbf{x}),$$

则 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数.

证明 对任意的 $(x, \eta) \in \text{epi} f$, 有 $\eta \geq f(x) = \sup_{i \in I} f_i(x)$, 它等价于

$$\eta \geq f_i(x), i \in I.$$

这意味着

$$(x, \eta) \in \text{epi} f_i, i \in I,$$

或即 $(x, \eta) \in \bigcap_{i \in I} \text{epi} f_i$, 因此 $\text{epi} f = \bigcap_{i \in I} \text{epi} f_i$. 由于 f_i ($i \in I$) 是凸函数, 据定理 2.1.5 知 $\text{epi} f_i$ ($i \in I$) 是凸集, 再由定理 1.1.2 的 (2), 得 $\text{epi} f = \bigcap_{i \in I} \text{epi} f_i$ 是凸集. 于是, 由定理 2.1.5 即知 f 是凸函数. \square

注 2.1.6 特别是, 在定理 2.1.7 中设各 f_i ($i \in I$) 是仿射函数, 则该定理表示线性空间上仿射函数的上包络是凸函数. 事实上, 能够证明线性空间上的正常凸函数在其有效域内部可以用仿射函数的上包络表示, 而线性拓扑空间上的下半连续正常凸函数可以用仿射函数的上包络表示^[24].

定理 2.1.8 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 和 $g: R \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 并且 $g(t) = f(x + td)$ (其中 $d \in \mathcal{X}$), 则 f 是凸函数当且仅当对任意的 $d \in \mathcal{X}$, g 是凸函数.

证明 必要性. 对任意的 $d \in \mathcal{X}$, 任取 $t_1, t_2 \in R$ 和 $\lambda \in (0, 1)$, 由于 f 是凸函数, 有

$$\begin{aligned} g(\lambda t_1 + (1 - \lambda)t_2) &= f(x + [\lambda t_1 + (1 - \lambda)t_2]d) \\ &= f(\lambda[x + t_1 d] + (1 - \lambda)[x + t_2 d]) \\ &\leq \lambda f(x + t_1 d) + (1 - \lambda)f(x + t_2 d) \\ &= \lambda g(t_1) + (1 - \lambda)g(t_2). \end{aligned}$$

因此, g 是凸函数.

充分性. 任取 $x^1, x^2 \in \mathcal{X}$ 和 $\lambda \in (0, 1)$, 则有

$$f(\lambda x^1 + (1-\lambda)x^2) = f(x^1 + [\lambda \cdot 0 + (1-\lambda) \cdot 1](x^2 - x^1)).$$

令 $x^1 = x$, $\lambda \cdot 0 + (1-\lambda) \cdot 1 = t$, $x^2 - x^1 = d$, 从上式注意到 g 是凸函数, 可得

$$\begin{aligned} f(\lambda x^1 + (1-\lambda)x^2) &= g(\lambda \cdot 0 + (1-\lambda) \cdot 1) \\ &\leq \lambda g(0) + (1-\lambda)g(1) \\ &= \lambda f(x^1) + (1-\lambda)f(x^2), \end{aligned}$$

从而 f 是凸函数. \square

定理 2.1.9 设 $h: \mathcal{X} \rightarrow R$ 是凸函数, $g: R \rightarrow R \cup \{\pm\infty\}$ 是非减的正常凸函数, $f(x) = g(h(x))$, 则 $f: \mathcal{X} \rightarrow R \cup \{\pm\infty\}$ 是凸函数.

证明 任取 $x^1, x^2 \in \mathcal{X}$ 和 $\lambda \in (0, 1)$, 由于 h 是凸的, g 是非减的, 从而有

$$\begin{aligned} f(\lambda x^1 + (1-\lambda)x^2) &= g(h(\lambda x^1 + (1-\lambda)x^2)) \\ &\leq g(\lambda h(x^1) + (1-\lambda)h(x^2)). \end{aligned} \tag{2.1.8}$$

又因 g 是正常凸函数, 故有

$$\begin{aligned} g(\lambda h(x^1) + (1-\lambda)h(x^2)) &\leq \lambda g(h(x^1)) + (1-\lambda)g(h(x^2)) \\ &= \lambda f(x^1) + (1-\lambda)f(x^2). \end{aligned}$$

将上式代入 (2.1.8) 得到

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda)f(x^2),$$

故 f 是凸函数. \square

利用函数上图象的凸包, 可以给出函数的凸包概念.

定义 2.1.7 设 $f: \mathcal{X} \rightarrow R \cup \{\pm\infty\}$ 是广义实值函数, 令

$$g(x) = \inf\{\eta \mid (x, \eta) \in \text{co}(\text{epi} f)\},$$

则称 $g: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是 f 的凸包(函数), 记作 $\text{co}(f)$.

为刻画凸包函数的特性, 再引进函数的弱函数和弱凸函数概念如下.

定义 2.1.8 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 和 $h: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数. 若

$$h(x) \leq f(x) \quad \forall x \in \mathcal{X},$$

则称 h 是 f 的弱函数. 若 h 还是 \mathcal{X} 上的凸函数, 则称 h 是 f 的弱凸函数.

下述定理表明, 任意函数的凸包是一凸函数.

定理 2.1.10 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 则 $\text{co}(f)$ 是 f 的最大弱凸函数.

证明 由 $\text{co}(\text{epi} f)$ 是凸集, 从定义 2.1.7 可以推知 $\text{co}(f)$ 是凸函数, 并且有

$$\text{epi} f \subset \text{co}(\text{epi} f) \subset \text{epi}(\text{co}(f)).$$

据此, 由定义 2.1.6 的(1)得

$$\text{co}(f)(x) \leq f(x) \quad \forall x \in \mathcal{X},$$

因而按定义 2.1.8 知 $\text{co}(f)$ 是 f 的弱凸函数. 现设 $h: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是任一上图象包含 $\text{co}(\text{epi} f)$ 的凸函数. 由于

$$h(x) = \inf \{ \eta : h(x) \leq \eta \} = \inf \{ \eta \mid h(x) \in \text{epi} h \},$$

$$\text{co}(f)(x) = \inf \{ \eta \mid h(x) \in \text{epi} f \},$$

而 $\text{co}(\text{epi} f) \subset \text{epi} h$, 故有

$$h(x) \leq \text{co}(f)(x) \quad \forall x \in \mathcal{X}.$$

由 h 的任意性, 即得 $\text{co}(f)$ 是 f 的最大弱凸函数. \square

最后, 对于 Euclid 空间 R^n 上的可微凸函数, 我们依次给出它的一阶和二阶的判别条件.

设 $f: R^n \rightarrow R$ 一阶可微, $\mathbf{x} = (x_1, \dots, x_n)^T$, 记

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^T$$

是 f 在点 \mathbf{x} 处的梯度.

定理 2.1.11 设 $S \subset R^n$ 是开凸集, $f: S \rightarrow R$ 一阶可微.

(1) f 是 S 上的凸函数当且仅当

$$\nabla f(\mathbf{x}^2)^T(\mathbf{x}^1 - \mathbf{x}^2) \leq f(\mathbf{x}^1) - f(\mathbf{x}^2) \quad \forall \mathbf{x}^1, \mathbf{x}^2 \in S. \quad (2.1.9)$$

(2) f 是 S 上的严格(强)凸函数当且仅当

$$\begin{aligned} \nabla f(\mathbf{x}^2)^T(\mathbf{x}^1 - \mathbf{x}^2) &< f(\mathbf{x}^1) - f(\mathbf{x}^2) \quad \forall \mathbf{x}^1, \mathbf{x}^2 \in S, \\ f(\mathbf{x}^1) &\neq f(\mathbf{x}^2) \quad (\mathbf{x}^1 \neq \mathbf{x}^2). \end{aligned} \quad (2.1.10)$$

证明 (1) 必要性. 从 f 是 S 上的凸函数, 由定义 2.1.1 的 (1) 知, 对任意的 $\mathbf{x}^1, \mathbf{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 有

$$f(\lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2) \leq \lambda f(\mathbf{x}^1) + (1 - \lambda)f(\mathbf{x}^2),$$

从而

$$\frac{f(\mathbf{x}^2 + \lambda(\mathbf{x}^1 - \mathbf{x}^2)) - f(\mathbf{x}^2)}{\lambda} \leq f(\mathbf{x}^1) - f(\mathbf{x}^2).$$

将上式左端的分子作 Taylor 展开, 则有

$$\nabla f(\mathbf{x}^2)^T(\mathbf{x}^1 - \mathbf{x}^2) + \frac{o(\|\lambda(\mathbf{x}^1 - \mathbf{x}^2)\|)}{\lambda} \leq f(\mathbf{x}^1) - f(\mathbf{x}^2),$$

令 $\lambda \rightarrow 0$ 便得 (2.1.9).

充分性. 对任意的 $\mathbf{x}^1, \mathbf{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 因为 S 是凸集, 故由定义 1.1.1 知

$$\mathbf{x} = \lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2 \in S. \quad (2.1.11)$$

在 (2.1.9) 中令 $\mathbf{x}^1 = \mathbf{x}^1$, $\mathbf{x}^2 = \mathbf{x}$, 则有

$$\nabla f(\mathbf{x})^T(\mathbf{x}^1 - \mathbf{x}) \leq f(\mathbf{x}^1) - f(\mathbf{x}) \quad \forall \mathbf{x}^1 \in S, \quad (2.1.12)$$

再在(2.1.9)中令 $\mathbf{x}^1 = \mathbf{x}^2$, $\mathbf{x}^2 = \mathbf{x}$, 又有

$$\nabla f(\mathbf{x})^T (\mathbf{x}^2 - \mathbf{x}) \leq f(\mathbf{x}^2) - f(\mathbf{x}) \quad \forall \mathbf{x}^2 \in S. \quad (2.1.13)$$

作 $\lambda(2.1.12) + (1 - \lambda)(2.1.13)$, 得到

$$\begin{aligned} & \nabla f(\mathbf{x})^T [\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2 - \mathbf{x}] \\ & \leq \lambda f(\mathbf{x}^1) + (1 - \lambda) f(\mathbf{x}^2) - f(\mathbf{x}). \end{aligned}$$

注意到(2.1.11), 便有

$$f(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) = f(\mathbf{x}) \leq \lambda f(\mathbf{x}^1) + (1 - \lambda) f(\mathbf{x}^2),$$

于是由定义 2.1.1 的(1)得 f 是 S 上的凸函数.

(2) 将(1)证明中的不等式改为严格不等式, 类似地可以推证. \square

设 $f: R^n \rightarrow R$ 二阶可微, $\mathbf{x} = (x_1, \dots, x_n)^T$, 记

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}$$

是 f 在点 \mathbf{x} 处的 Hesse 矩阵.

定理 2.1.12 设 $S \subset R^n$ 是开凸集, $f: S \rightarrow R$ 二阶连续可微.

(1) f 是 S 上的凸函数当且仅当对任意的 $\mathbf{x} \in S$, $\nabla^2 f(\mathbf{x})$ 是半正定的.

(2) 若对任意的 $\mathbf{x} \in S$, $\nabla^2 f(\mathbf{x})$ 是正定的, 则 f 是 S 上的强凸函数.

证明 (1) 充分性. 对任意的 $\mathbf{x}^1, \mathbf{x}^2 \in S$, 有 Taylor 公式

$$\begin{aligned} f(\mathbf{x}^1) &= f(\mathbf{x}^2) + \nabla f(\mathbf{x}^2)^T (\mathbf{x}^1 - \mathbf{x}^2) \\ &\quad + \frac{1}{2} (\mathbf{x}^1 - \mathbf{x}^2)^T \nabla^2 f(\xi) (\mathbf{x}^1 - \mathbf{x}^2), \end{aligned}$$

其中 $\xi \in S$. 由于 $\nabla^2 f(\xi)$ 是半正定的, 故 $(x^1 - x^2)^T \nabla^2 f(\xi) (x^1 - x^2) \geq 0$, 于是从上式推知

$$\nabla f(x^2)^T (x^1 - x^2) \leq f(x^1) - f(x^2).$$

据此, 由定理 2.1.11 的(1)得知 f 在 S 上是凸的.

必要性. 任取 $x \in S$, 因为 S 是开集, 故对任意的 $y \in R^n$, 存在充分小的 $\lambda > 0$ 有 $x + \lambda y \in S$. 由于 f 是 S 上的凸函数, 在定理 2.1.11 中的(1)中令 $x^1 = x + \lambda y$, $x^2 = x$, 则得

$$f(x) + \lambda \nabla f(x)^T y \leq f(x + \lambda y). \quad (2.1.14)$$

另外, 由 Taylor 公式有

$$f(x + \lambda y) = f(x) + \lambda \nabla f(x)^T y + \frac{1}{2} \lambda^2 y^T \nabla^2 f(\xi) y,$$

其中 $\xi = x + \theta \lambda y$ ($\theta \in (0, 1)$). 将它与(2.1.14)作比较, 得知 $y^T \nabla^2 f(\xi) y \geq 0$. 令 $\lambda \rightarrow 0$, 则 $\xi \rightarrow x$, 于是得到 $y^T \nabla^2 f(x) y \geq 0$, 即 $\nabla^2 f(x)$ 是半正定的.

(2) 对任意的 $x^1, x^2 \in S$, $x^1 \neq x^2$, 由 Taylor 公式有

$$\begin{aligned} f(x^1) &= f(x^2) + \nabla f(x^2)^T (x^1 - x^2) \\ &\quad + \frac{1}{2} (x^1 - x^2)^T \nabla^2 f(\xi) (x^1 - x^2), \end{aligned} \quad (2.1.15)$$

其中 $\xi \in S$. 因对任意的 $x \in S$, $\nabla^2 f(x)$ 是正定的, 故 $(x^1 - x^2)^T \nabla^2 f(\xi) (x^1 - x^2) > 0$. 注意到 $x^1 \neq x^2$, 从(2.1.15)推知

$$\nabla f(x^2)^T (x^1 - x^2) < f(x^1) - f(x^2).$$

据此, 由定理 2.1.11 的(2)得到 f 在 S 上是强凸的. \square

§ 2.2 凸函数的连续性和对偶性

连续性是函数的一个重要特性. 在这一节, 先讨论线性拓扑空

间上凸函数的连续性问题. 然后研究线性拓扑空间上的凸函数与对偶空间上相应的共轭函数之间的对偶关系.

设 \mathcal{X} 是线性拓扑空间.

定义 2.2.1 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 点 $x^0 \in \mathcal{X}$. 若对任意的 $\varepsilon > 0$, 存在点 x^0 的邻域 $U(x^0)$, 有

$$|f(x) - f(x^0)| < \varepsilon \quad \forall x \in U(x^0),$$

则称 f 在点 x^0 处是连续的. 若 f 在集合 $S \subset \mathcal{X}$ 的每一点处都是连续的, 则称 f 是 S 上的连续函数(连续泛函), 或 f 在 S 上是连续的.

注 2.2.1 由定义 2.2.1 和定义 2.1.2 易知, 若 f 在点 $x^0 \in \mathcal{X}$ 处是连续的, 则 f 在点 x^0 处既是下半连续的又是上半连续的; 反之, 若 f 在点 x^0 处既是下半连续的又是上半连续的, 则 f 在点 x^0 处是连续的.

定理 2.2.1 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数. 若存在点 $x^0 \in \text{int}(\text{dom } f)$ 使得 f 在 x^0 的某邻域内上有界, 则 f 在 $\text{int}(\text{dom } f)$ 内是连续的.

证明 不失一般性, 我们设 $x^0 = \theta$ (否则, 可令 $y = x - x^0$ 将 $f(x)$ 转为对 $g(y) = f(y + x^0)$ 的讨论). 由已设 f 在 θ 的邻域内上有界, 则存在点 θ 的对称邻域 $U(\theta)$ 和常数 $M \in R$, 有

$$f(x) \leq M \quad \forall x \in U(\theta). \quad (2.2.1)$$

先证 f 在点 θ 处是连续的. 为此, 任给 $\varepsilon \in (0, 1)$, 则当 $x \in \varepsilon U(\theta)$ 时有 $\pm \frac{x}{\varepsilon} \in U(\theta)$, 从而根据 f 是 $U(\theta)$ 上的凸函数, 按定义 2.1.1 的(1)有

$$f(x) = f\left(\varepsilon \frac{x}{\varepsilon} + (1 - \varepsilon) \cdot \theta\right) \leq \varepsilon f\left(\frac{x}{\varepsilon}\right) + (1 - \varepsilon)f(\theta).$$

注意到(2.2.1), 有 $f(x) \leq \varepsilon M + (1 - \varepsilon)f(\theta)$, 故得

$$f(x) - f(\theta) \leq \varepsilon [M - f(\theta)]. \quad (2.2.2)$$

另外,由 f 是凸函数和(2.2.1)又有

$$\begin{aligned} f(\theta) &= f\left(\frac{1}{1+\epsilon}x + \frac{\epsilon}{1+\epsilon}\left(\frac{-x}{\epsilon}\right)\right) \\ &\leq \frac{1}{1+\epsilon}f(x) + \frac{\epsilon}{1+\epsilon}f\left(\frac{-x}{\epsilon}\right) \\ &\leq \frac{1}{1+\epsilon}f(x) + \frac{\epsilon}{1+\epsilon}M, \end{aligned}$$

于是

$$f(x) - f(\theta) \geq -\epsilon[M - f(\theta)]. \quad (2.2.3)$$

由(2.2.2)和(2.2.3)得到

$$|f(x) - f(\theta)| \leq \epsilon[M - f(\theta)],$$

因为 $\epsilon > 0$ 是任给的,所以由定义 2.2.1 知 f 在点 θ 处是连续的.

现任取 $x \in \text{int}(\text{dom} f)$, 我们证明 f 在点 x 处是连续的. 事实上,这时存在 $\rho > 1$ 使得 $\rho x \in \text{dom} f$. 因为 $U(\theta)$ 是点 θ 的邻域,故

$$U(x) = x + \left(1 - \frac{1}{\rho}\right)U(\theta)$$

是点 x 的邻域. 由此,对任意的 $y \in U(x)$, 存在 $u \in U(\theta)$ 有

$$y = x + \left(1 - \frac{1}{\rho}\right)u.$$

根据 f 在 $\text{dom} f$ 上是凸的以及(2.2.1),我们得到

$$\begin{aligned} f(y) &= f\left(\frac{1}{\rho}(\rho x) + \left(1 - \frac{1}{\rho}\right)u\right) \\ &\leq \frac{1}{\rho}f(\rho x) + \left(1 - \frac{1}{\rho}\right)f(u) \\ &\leq \frac{1}{\rho}f(\rho x) + \left(1 - \frac{1}{\rho}\right)M. \end{aligned}$$

由 $y \in U(x)$ 的任意性,即知 f 在点 x 的邻域上有界. 于是,与前面

已证得的结果同理,可得 f 在点 x 处是连续的. \square

推论 2.2.2 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数,则 f 在 $\text{int}(\text{dom} f)$ 内是连续的当且仅当 $\text{int}(\text{epi} f) \neq \emptyset$.

证明 因为 $\text{int}(\text{epi} f) \neq \emptyset$ 等价于 f 在某些点的邻域内上有界,所以由定理 2.2.1 并注意到连续函数总是上有界的,它也等价于 f 在 $\text{int}(\text{dom} f)$ 内是连续的. \square

对于有限维 Euclid 空间 R^n 上的正常凸函数,则由定理 2.2.1 可以推知它在有效域内部一定是连续的.

推论 2.2.3 设 $f: R^n \rightarrow R \cup \{+\infty\}$ 是正常凸函数,则 f 在 $\text{int}(\text{dom} f)$ 内是连续的.

证明 在 $\text{int}(\text{dom} f)$ 内可以找到 $n+1$ 个仿射无关的点 x^0, x^1, \dots, x^n , 以它们为顶点构成的单纯形 $A \subset \text{int}(\text{dom} f)$, 并且由 f 的凸性有

$$f(x) = f\left(\sum_{i=0}^n \lambda_i x^i\right) \leq \sum_{i=0}^n \lambda_i f(x^i) \quad \forall x = \sum_{i=0}^n \lambda_i x^i \in A,$$

其中 $\lambda_i \geq 0$ ($i = 0, 1, \dots, n$), $\sum_{i=0}^n \lambda_i = 1$. 据此, f 在 A 的内部上有界,而 A 的内部是 R^n 中的开集,故由定理 2.2.1 得证. \square

但是,对于无限维空间来说,上述结论一般不成立.事实上,在线性拓扑空间上可以存在处处有限而处处不连续的凸函数.

下面考虑半连续情况.

定理 2.2.4 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数.若存在点 $x^0 \in \text{dom} f$ 使得 f 在 x^0 处是上半连续的,则 f 在 $\text{int}(\text{dom} f)$ 内是连续的.

证明 因为 f 在点 x^0 处是上半连续的,由定义 2.1.2 的(2)可知 f 在 x^0 的某邻域内上有界.于是,根据定理 2.2.1 即得结论. \square

函数的半连续性与相应集合的闭性之间有如下关系.

定理 2.2.5 设 $f: \mathcal{X} \rightarrow R \cup \{\pm\infty\}$ 是广义实值函数,则以下条件是等价的:

- (1) f 是下半连续函数;
 (2) $\text{epi} f \subset \mathcal{X} \times R$ 是闭集;
 (3) 对任一 $c \in R$, $\{x \in \mathcal{X} \mid f(x) \leq c\}$ 是闭集;
 (4) 对任一 $c \in R$, $\{x \in \mathcal{X} \mid f(x) > c\}$ 是开集.

证明 (1) \Rightarrow (2). 任取点列 $\{(x^k, \eta_k)\} \subset \text{epi} f$, 并且 $(x^k, \eta_k) \rightarrow (x^0, \eta_0)$ ($k \rightarrow \infty$). 因为 $(x^k, \eta_k) \in \text{epi} f$ ($k = 1, 2, \dots$), 所以由定义 2.1.6 的(1)有

$$f(x^k) \leq \eta_k, k = 1, 2, \dots.$$

由于 f 在点 x^0 处是下半连续的, 于是由定义 2.1.2 的(1)可知, 对任给的 $\varepsilon > 0$, 存在 x^0 的邻域 $U(x^0)$ 有

$$f(x^0) - \varepsilon \leq \inf_{x \in U(x^0)} f(x).$$

令 $\varepsilon \rightarrow 0$, 特别地有

$$f(x^0) \leq \liminf_{k \rightarrow \infty} f(x^k) \leq \liminf_{k \rightarrow \infty} \eta_k = \eta_0.$$

据此, 再按定义 2.1.6 的(1)即得 $(x^0, \eta_0) \in \text{epi} f$, 故知 $\text{epi} f$ 是闭集.

(2) \Rightarrow (3). 任取点列 $\{x^k\} \subset \{x \in \mathcal{X} \mid f(x) \leq c\}$, 并且 $x^k \rightarrow x^0$ ($k \rightarrow \infty$). 因为 $f(x^k) \leq c$ ($k = 1, 2, \dots$), 由定义 2.1.6 的(1)有 $(x^k, c) \in \text{epi} f$. 由于 $\text{epi} f$ 是闭集, 因此 $f(x^0) \leq c$. 于是有 $x^0 \in \{x \in \mathcal{X} \mid f(x) \leq c\}$, 从而得证.

(3) \Rightarrow (4). 由 $\{x \in \mathcal{X} \mid f(x) \leq c\}$ 是闭集, 即知 $\{x \in \mathcal{X} \mid f(x) > c\} = \mathcal{X} \setminus \{x \in \mathcal{X} \mid f(x) \leq c\}$ 是开集.

(4) \Rightarrow (1). 任取 $x^0 \in \mathcal{X}$, 对任给的 $\varepsilon > 0$ 令

$$c = f(x^0) - \varepsilon, \quad (2.2.4)$$

则 $f(x^0) = c + \varepsilon > c$, 于是 $x^0 \in \{x \in \mathcal{X} \mid f(x) > c\}$. 因为 $\{x \in \mathcal{X} \mid f(x) > c\}$ 是开集, 故存在点 x^0 的邻域 $U(x^0)$, 有 $U(x^0) \subset \{x \in \mathcal{X} \mid f(x) > c\}$. 据此, 有

$$f(x) > c \quad \forall x \in U(x^0),$$

或由(2.2.4)得

$$-\varepsilon < f(x) - f(x^0) \quad \forall x \in U(x^0).$$

按定义 2.1.2 的(1), 即知 f 在 x^0 处是下半连续的. 因为 $x^0 \in \mathcal{X}$ 是任取的, 故 f 是下半连续函数. \square

定理 2.2.6 设 $f: \mathcal{X} \rightarrow R \cup \{\pm\infty\}$ 是广义实值函数, 则以下条件是等价的:

- (1) f 是上半连续函数;
- (2) $\text{hpy} f \subset \mathcal{X} \times R$ 是闭集;
- (3) 对任一 $c \in R$, $\{x \in \mathcal{X} \mid f(x) \geq c\}$ 是闭集;
- (4) 对任一 $c \in R$, $\{x \in \mathcal{X} \mid f(x) < c\}$ 是开集.

证明 与定理 2.2.5 的证明类似, 在此从略. \square

定理 2.2.7 设 \mathcal{B} 是 Banach 空间, $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常凸函数. 若 f 在 \mathcal{B} 上是下半连续的, 则 f 在 $\text{int}(\text{dom} f)$ 内是连续的.

证明 不妨取 $\theta \in \text{int}(\text{dom} f)$, 并且 $f(\theta) < c$. 记

$$A = \{x \in \mathcal{B} \mid f(x) \leq c\},$$

因为 f 在 \mathcal{B} 上是下半连续的, 由定理 2.2.5 的(1)和(3)知 A 是闭的. 又由 f 的凸性, 根据定理 2.1.3 知 A 是凸集. 对任意的 $x \in \mathcal{B}$, 因为 f 是 \mathcal{B} 上的凸函数, 故存在 $\bar{\varepsilon} > 0$, 对任意充分小的 $\varepsilon \in (0, \bar{\varepsilon})$ 有

$$f\left(\frac{\varepsilon x}{1+\varepsilon}\right) \leq \frac{1}{1+\varepsilon}f(\theta) + \frac{\varepsilon}{1+\varepsilon}f(x) \leq c.$$

由此, 有

$$\frac{\varepsilon}{1+\varepsilon}x \in A, \quad \text{或} \quad x \in \frac{1+\varepsilon}{\varepsilon}A.$$

这说明 A 是吸收集, 从而 $U = A \cap (-A)$ 是对称吸收闭凸集. 现在证明 $\text{int} U \neq \emptyset$. 事实上, 假若 $\text{int} U = \emptyset$, 则有

$$\text{int}(kU) = \emptyset, k = 1, 2, \dots \quad (2.2.5)$$

由(2.2.5), 当 $k = 1$ 时知, 任取 $\mathbf{x}^1 \notin U$, 则存在 $\delta_1 > 0$, 使对 \mathbf{x}^1 的 δ_1 -邻域 $U(\mathbf{x}^1, \delta_1)$ 有

$$U \cap U(\mathbf{x}^1, \delta_1) = \emptyset.$$

由(2.2.5), 当 $k = 2$ 时, 又知 $U(\mathbf{x}^1, \delta_1) \not\subset 2U$. 于是, 取 $\mathbf{x}^2 \in U(\mathbf{x}^1, \delta_1)$ 和 $\mathbf{x}^2 \notin 2U$, 则存在 $\delta_2 \in \left(0, \frac{\delta_1}{2}\right]$, 使对 \mathbf{x}^2 的 δ_2 -邻域 $U(\mathbf{x}^2, \delta_2)$ 有

$$2U \cap U(\mathbf{x}^2, \delta_2) = \emptyset.$$

同样地, 由(2.2.5), 当 $k = 3$ 时, 则知 $U(\mathbf{x}^2, \delta_2) \not\subset 3U$, 故取 $\mathbf{x}^3 \in U(\mathbf{x}^2, \delta_2)$ 和 $\mathbf{x}^3 \notin 3U$, 存在 $\delta_3 \in \left(0, \frac{\delta_2}{2}\right]$, 使对 \mathbf{x}^3 的 δ_3 -邻域 $U(\mathbf{x}^3, \delta_3)$ 有

$$3U \cap U(\mathbf{x}^3, \delta_3) = \emptyset.$$

如此继续下去, 可得到序列 $\{\mathbf{x}^k\}$ 和正数列 $\{\delta_k\}$, 满足 $\mathbf{x}^{k+1} \in U(\mathbf{x}^k, \delta_k)$ 和 $\delta_{k+1} \in \left(0, \frac{\delta_k}{2}\right]$ ($k = 1, 2, \dots$), 并且有

$$kU \cap U(\mathbf{x}^k, \delta_k) = \emptyset. \quad (2.2.6)$$

显然, $\{\mathbf{x}^k\}$ 是 Cauchy 序列, 由于 \mathcal{B} 是完备空间, 则有 $\mathbf{x}^k \rightarrow \mathbf{x}^* \in \mathcal{B}$. 据此, 由(2.2.6)得知 $\mathbf{x}^* \notin kU$ ($k = 1, 2, \dots$), 这与 U 是吸收集矛盾.

从 $\text{int}U \neq \emptyset$, 设 $\mathbf{x}^0 \in \text{int}U$, 则存在 $\delta > 0$ 对于 \mathbf{x}^0 的 δ -邻域 $U(\mathbf{x}^0, \delta)$ 有 $U(\mathbf{x}^0, \delta) \subset U$. 因为 U 是对称凸集, 故有 $-\mathbf{x}^0 \in U$. 设 $U(\mathbf{0}, \delta)$ 是点 $\mathbf{0}$ 的 δ -邻域, 则有

$$\frac{1}{2}U(\mathbf{0}, \delta) = \frac{1}{2}[U(\mathbf{x}^0, \delta) - \mathbf{x}^0] \subset \frac{1}{2}U + \frac{1}{2}U = U.$$

显然, f 在 $\frac{1}{2}U(\mathbf{0}, \delta)$ 内是上有界的, 于是 f 在点 $\mathbf{0}$ 的邻域 U 内上

有界. 据此, 由定理 2.2.1 得知 f 在 $\text{int}(\text{dom} f)$ 内是连续的. \square

定义 2.2.2 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$, $g: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数. 若

$$\text{epi} g = \text{cl}(\text{epi} f),$$

则称 g 是 f 的下半连续包.

定义 2.2.3 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数.

(1) 若 $g: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是 f 的下半连续包, 则称 g 是 f 的闭包(函数), 记作 $g = \text{cl}(f)$.

(2) 若 $f = \text{cl}(f)$, 则称 f 是闭凸函数, 或 f 是闭凸的.

注 2.2.2 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, 则由定义 2.2.2 和定义 2.2.3 的(1)可知

$$\text{epi}(\text{cl}(f)) = \text{cl}(\text{epi} f). \quad (2.2.7)$$

定理 2.2.8 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, 则 f 是 \mathcal{X} 上的下半连续函数当且仅当 f 是闭凸函数.

证明 充分性. 从 f 是闭凸函数, 由定义 2.2.3 的(2)有 $f = \text{cl}(f)$. 据此, 由(2.2.7)可知 $\text{epi} f = \text{epi}(\text{cl}(f)) = \text{cl}(\text{epi} f)$, 因而 $\text{epi} f$ 是闭集. 于是, 由定理 2.2.5 的(2)和(1), 得知 f 是下半连续的.

必要性. 设 f 是下半连续的, 则由定理 2.2.5 知 $\text{epi} f$ 是闭集. 因此, 由(2.2.7)有 $\text{epi} f = \text{cl}(\text{epi} f) = \text{epi}(\text{cl}(f))$. 这说明 $f = \text{cl}(f)$, 从而由定义 2.2.3 的(2)得 f 是闭凸函数. \square

以下建立线性拓扑空间上的凸函数与对偶空间上相应的共轭函数间的对偶关系.

定义 2.2.4 设 $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凸函数, \mathcal{X}^* 是 \mathcal{X} 的对偶空间.

(1) 令

$$f^*(x^*) = \sup_{x \in \mathcal{X}} \{\langle x^*, x \rangle - f(x)\}, \quad x^* \in \mathcal{X}^*, \quad (2.2.8)$$

则称 $f^*: \mathcal{X}^* \rightarrow R \cup \{\pm \infty\}$ 是 f 的共轭函数.

(2) 令

$$f^{**}(x) = \sup_{x^* \in \mathcal{X}^*} \{\langle x^*, x \rangle - f^*(x^*)\}, \quad x \in \mathcal{X}, \quad (2.2.9)$$

则称 $f^{**}: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是 f 的二次共轭函数.

通常, 称由 (2.2.8) 所确定的 f 与 f^* 之间的关系具有共轭关系或对偶关系. 显然, 由定义 2.2.4, 即知有 $(f^*)^* = f^{**}$.

注 2.2.3 从定义 2.2.4 的 (1), 可以直接得到如下的 Young-Fenchel 不等式:

$$\langle x^*, x \rangle \leq f^*(x^*) + f(x) \quad \forall x^* \in \mathcal{X}^*, \forall x \in \mathcal{X}. \quad (2.2.10)$$

注 2.2.4 由定义 2.2.4 的 (1) 可知, 若存在某个 $\bar{x} \in \mathcal{X}$ 使得 $f(\bar{x}) = -\infty$, 则有 $f^*(x^*) \equiv +\infty$; 反之, 若存在某个 $\bar{x}^* \in \mathcal{X}^*$ 使得 $f^*(\bar{x}^*) = -\infty$, 则仅当 $f(x) \equiv +\infty$ 时才有可能, 从而使 $f^*(x^*) \equiv -\infty$. 因此, 当 $f \equiv +\infty$ 或 f 取 $-\infty$ 的情形都是没有意义的. 此后, 我们只考虑 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 并且 $f \not\equiv +\infty$ 的情形.

以下介绍共轭函数的有关性质.

定理 2.2.9 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是凸函数, 则

$$\inf_{x \in \mathcal{X}} f(x) = -f^*(\theta).$$

证明 由定义 2.2.4 的 (1) 直接可知 $f^*(\theta) = \sup_{x \in \mathcal{X}} \{\langle \theta, x \rangle - f(x)\} = -\inf_{x \in \mathcal{X}} f(x)$. \square

定理 2.2.10 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是凸函数.

(1) $f^*: \mathcal{X}^* \rightarrow R \cup \{+\infty\}$ 是下半连续正常凸函数.

(2) $f^{**}: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是下半连续正常凸函数.

证明 (1) 为证 f^* 是下半连续的, 由定理 2.2.5, 我们证明 $\text{epi} f^*$ 是闭集. 事实上, 取点列 $\{(x^k, \eta_k)\} \subset \text{epi} f^*$, 并且 $(x^k, \eta_k) \rightarrow (x^0, \eta_0)$ ($k \rightarrow \infty$), 因为 $(x^k, \eta_k) \in \text{epi} f^*$ ($k = 1, 2, \dots$), 所以由定义 2.1.6 的 (1) 和定义 2.2.4 的 (1) 有

$$\eta_k \geq f^*(x^k) = \sup_{x \in \mathcal{X}} \{ \langle x^k, x \rangle - f(x) \}, \quad x^k \in \mathcal{X}^*, \quad k=1, 2, \dots$$

令 $k \rightarrow \infty$, 则得

$$\begin{aligned} \eta_0 &= \lim_{k \rightarrow \infty} \eta_k \geq \lim_{k \rightarrow \infty} \sup_{x \in \mathcal{X}} \{ \langle x^k, x \rangle - f(x) \} \\ &= \sup_{x \in \mathcal{X}} \{ \langle x^0, x \rangle - f(x) \} = f^*(x^0). \end{aligned}$$

由此, 我们有 $(x^0, \eta_0) \in \text{epi} f^*$, 故 $\text{epi} f^*$ 是闭的.

现证 f^* 是凸的. 为此, 任取 $x^1, x^2 \in \mathcal{X}^*$ 和 $\lambda \in (0, 1)$, 由定义 2.2.4 的(1)有

$$\begin{aligned} f^*(\lambda x^1 + (1-\lambda)x^2) &= \sup_{x \in \mathcal{X}} \{ \langle \lambda x^1 + (1-\lambda)x^2, x \rangle - f(x) \} \\ &\leq \lambda \sup_{x \in \mathcal{X}} \{ \langle x^1, x \rangle - f(x) \} \\ &\quad + (1-\lambda) \sup_{x \in \mathcal{X}} \{ \langle x^2, x \rangle - f(x) \} \\ &= \lambda f^*(x^1) + (1-\lambda)f^*(x^2). \end{aligned}$$

于是, 根据定义 2.1.1 的(1)和定义 2.1.3, 即知 f^* 是凸函数. 由注 2.2.4, 因为 $f^* \neq -\infty$, 故 f^* 是正常凸函数.

(2) 与(1)的证明类似. \square

定理 2.2.11 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$, $h: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是凸函数. 若 h 是 f 的弱函数, 则 f^* 是 h^* 的弱函数.

证明 从 h 是 f 的弱函数, 由定义 2.1.8 有

$$h(x) \leq f(x) \quad \forall x \in \mathcal{X}.$$

据此, 按定义 2.2.4 的(1), 即知对任意的 $x^* \in \mathcal{X}^*$, 有

$$\begin{aligned} f^*(x^*) &= \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - f(x) \} \\ &\leq \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - h(x) \} \\ &= h^*(x^*), \end{aligned}$$

故 f^* 是 h^* 的弱函数. \square

定理 2.2.12 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是凸函数.

(1) f^{**} 是 f 的弱函数, 即对任意的 $x \in \mathcal{X}$ 有 $f^{**}(x) \leq f(x)$.

(2) $f^{**} = f$ 当且仅当 f 是下半连续正常凸函数.

(3) $f^{***} = f^*$.

证明 (1) 由 (2.2.10) (Young-Fenchel 不等式) 有

$$\langle x^*, x \rangle - f^*(x^*) \leq f(x) \quad \forall x^* \in \mathcal{X}^*, \forall x \in \mathcal{X}.$$

于是, 由定义 2.2.4 的 (2), 即得

$$f^{**}(x) = \sup_{x^* \in \mathcal{X}^*} \{\langle x^*, x \rangle - f^*(x^*)\} \leq f(x) \quad \forall x \in \mathcal{X}.$$

(2) 由定理 2.2.10 中的 (2) 知, $f = f^{**}$ 是下半连续正常凸函数. 反之, 设 f 是下半连续正常凸函数, 则按注 2.1.6, 它可以用仿射函数的上包络表示. 于是, 存在 $\{x_i^*\}_{i \in I} \subset \mathcal{X}^*$ 和 $\{\alpha_i\}_{i \in I} \in R$ (I 是指标集), 使得

$$f(x) = \sup_{i \in I} \{\langle x_i^*, x \rangle - \alpha_i\}. \quad (2.2.11)$$

据此, 由定义 2.2.4 的 (1) 知, 对 $x^* \in \mathcal{X}^*$ 有

$$\begin{aligned} f^*(x^*) &= \sup_{x \in \mathcal{X}} \left\{ \langle x^*, x \rangle - \sup_{i \in I} \{\langle x_i^*, x \rangle - \alpha_i\} \right\} \\ &= \inf_{i \in I} \sup_{x \in \mathcal{X}} \{\langle x^*, x \rangle - \langle x_i^*, x \rangle + \alpha_i\}. \end{aligned}$$

对于 f^* 和 f^{**} , 利用 (2.2.10) 可知, 对任意的 $x \in \mathcal{X}$ 和任意的 $i \in I$ 有

$$\begin{aligned} f^{**}(x) &\geq \langle x^*, x \rangle - f^*(x^*) \\ &= \langle x^*, x \rangle - \inf_{i \in I} \sup_{x \in \mathcal{X}} \{\langle x^*, x \rangle - \langle x_i^*, x \rangle + \alpha_i\} \\ &\geq \langle x^*, x \rangle - \sup_{x \in \mathcal{X}} \{\langle x^*, x \rangle - \langle x_i^*, x \rangle + \alpha_i\}. \end{aligned}$$

在最后的式子中取 $x^* = x_i^*$, 则有

$$f^{**}(x) \geq \langle x^*, x \rangle - \alpha_i \quad \forall i \in I.$$

再根据(2.2.11), 得到

$$f^{**}(x) \geq \sup_{i \in I} \{ \langle x_i^*, x \rangle - \alpha_i \} = f(x) \quad \forall x \in \mathcal{X}.$$

由已证明的(1), 即得 $f^{**} = f$.

(3) 由(1)知 f^{**} 是 f 的弱函数, 再据定理 2.2.11, 得 f^* 是 $(f^{**})^*$ 的弱函数, 即有

$$f^*(x^*) \leq f^{***}(x^*) \quad \forall x^* \in \mathcal{X}^*. \quad (2.2.12)$$

另外, 在定理 2.2.12 的(1)中用 f^* 替换 f , 则有

$$f^{***}(x^*) \leq f^*(x^*) \quad \forall x^* \in \mathcal{X}^*. \quad (2.2.13)$$

由(2.2.12)和(2.2.13), 便可得到结论. \square

定理 2.2.13 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是凸函数, $x \in \mathcal{X}$, $x^* \in \mathcal{X}^*$.

(1) 若 $a \in R$, $g(x) = f(x) + a$, 则 $g^*(x^*) = f^*(x^*) - a$.

(2) 若 $a > 0$, $g(x) = af(x)$, 则 $g^*(x^*) = af^*\left(\frac{x^*}{a}\right)$.

(3) 若 $y \in \mathcal{X}$, $g(x) = f(x + y)$, 则 $g^*(x^*) = f^*(x^*) - \langle x^*, y \rangle$.

(4) 若 $\alpha \neq 0$, $g(x) = f(\alpha x)$, 则 $g^*(x^*) = \frac{1}{\alpha} f^*(x^*)$.

证明 (1) 由定义 2.2.4 的(1), 对任意的 $x^* \in \mathcal{X}^*$ 有

$$\begin{aligned} g^*(x^*) &= \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - (f(x) + a) \} \\ &= \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - f(x) \} - a \\ &= f^*(x^*) - a. \end{aligned}$$

(2) 根据定义 2.2.4 的(1), 直接有

$$g^*(x^*) = \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - af(x) \}$$

$$\begin{aligned} &= a \sup_{x \in \mathcal{X}} \left\{ \left\langle \frac{x^*}{a}, x \right\rangle - f(x) \right\} \\ &= a f^* \left(\frac{x^*}{a} \right). \end{aligned}$$

(3) 按定义 2.2.4 的(1), 令 $z = x + y$, 则得

$$\begin{aligned} g^*(x^*) &= \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - f(x + y) \} \\ &= \sup_{z \in \mathcal{X}} \{ \langle x^*, z - y \rangle - f(z) \} \\ &= \sup_{z \in \mathcal{X}} \{ \langle x^*, z \rangle - f(z) \} - \langle x^*, y \rangle \\ &= f^*(x^*) - \langle x^*, y \rangle. \end{aligned}$$

(4) 由定义 2.2.4 的(1), 令 $z = \alpha x$, 即可得证. \square

定理 2.2.14 设 $f_i: \mathcal{X} \rightarrow R \cup \{+\infty\}$ ($i \in I$, I 是指标集) 是凸函数, \mathcal{X}^* 是 \mathcal{X} 的对偶空间.

(1) 对任意的 $x^* \in \mathcal{X}^*$, 有 $\left(\inf_{i \in I} f_i \right)^*(x^*) = \sup_{i \in I} f_i^*(x^*)$.

(2) 对任意的 $x^* \in \mathcal{X}^*$, 有 $\left(\sup_{i \in I} f_i \right)^*(x^*) \leq \inf_{i \in I} f_i^*(x^*)$.

证明 由定义 2.2.4 的(1), 对任意的 $x^* \in \mathcal{X}^*$ 有

$$\begin{aligned} \left(\inf_{i \in I} f_i \right)^*(x^*) &= \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - \inf_{i \in I} f_i(x) \} \\ &= \sup_{x \in \mathcal{X}} \left\{ \sup_{i \in I} [\langle x^*, x \rangle - f_i(x)] \right\} \\ &= \sup_{i \in I} \left[\sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - f_i(x) \} \right] \\ &= \sup_{i \in I} f_i^*(x^*). \end{aligned}$$

(2) 由定义 2.2.4 的(1), 有

$$\left(\sup_{i \in I} f_i \right)^*(x^*) = \sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - \sup_{i \in I} f_i(x) \}$$

$$\begin{aligned}
&= \sup_{x \in \mathcal{X}} \left\{ \inf_{i \in I} [\langle x^*, x \rangle - f_i(x)] \right\} \\
&\leq \inf_{i \in I} \left[\sup_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - f_i(x) \} \right] \\
&= \inf_{i \in I} f_i^*(x^*). \quad \square
\end{aligned}$$

最后,介绍一个基本的对偶定理.为此,要引进凹函数的共轭函数的概念.

定义 2.2.5 设 $g: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是凹函数, \mathcal{X}^* 是 \mathcal{X} 的对偶空间. 令

$$g^*(x^*) = \inf_{x \in \mathcal{X}} \{ \langle x^*, x \rangle - g(x) \},$$

则称 $g^*: \mathcal{X}^* \rightarrow R \cup \{\pm \infty\}$ 是 g 的共轭函数.

与注 2.2.4 同样的理由,只需考虑 $g: \mathcal{X} \rightarrow R \cup \{-\infty\}$ 并且 $g \neq -\infty$ 的情形.

注 2.2.5 由定义 2.2.5 可以推知,对于凹函数的共轭函数,有与(2.2.10)类似的如下不等式:

$$g^*(x^*) + g(x) \leq \langle x^*, x \rangle \quad \forall x^* \in \mathcal{X}^*, \forall x \in \mathcal{X}. \quad (2.2.14)$$

关于凸函数和凹函数的共轭函数之间,有下面的对偶关系.

定理 2.2.15 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $g: \mathcal{X} \rightarrow R \cup \{-\infty\}$ 是正常凹函数. 若 f 或 g 在某点 $x^0 \in \text{dom} f \cap \text{dom} g$ 处是连续的, 则

$$\inf_{x \in \mathcal{X}} \{ f(x) - g(x) \} = \sup_{x^* \in \mathcal{X}^*} \{ g^*(x^*) - f^*(x^*) \}.$$

证明 由(2.2.14)和(2.2.10),可知

$$\begin{aligned}
g^*(x^*) + g(x) &\leq \langle x^*, x \rangle \leq f^*(x^*) + f(x) \\
&\quad \forall x^* \in \mathcal{X}^*, \forall x \in \mathcal{X}.
\end{aligned}$$

据此,有

$$f(x) - g(x) \geq g^*(x^*) - f^*(x^*)$$

$$\forall x^* \in \mathcal{X}^*, \forall x \in \mathcal{X},$$

从而得

$$\inf_{x \in \mathcal{X}} \{f(x) - g(x)\} \geq \sup_{x^* \in \mathcal{X}^*} \{g^*(x^*) - f^*(x^*)\}. \quad (2.2.15)$$

我们记

$$m = \inf_{x \in \mathcal{X}} \{f(x) - g(x)\}. \quad (2.2.16)$$

若 $m = -\infty$, 则由(2.2.15), 定理的结论成立. 设 $m > -\infty$, 作集合

$$S_1 = \{(x, \eta) \in \mathcal{X} \times R \mid g(x) + m \geq \eta\}, \quad (2.2.17)$$

$$S_2 = \text{epi} f = \{(x, \eta) \in \mathcal{X} \times R \mid \eta \geq f(x)\}. \quad (2.2.18)$$

从 $f \not\equiv +\infty$ 是凸函数, 由定理 2.1.5 知 S_2 是凸集, 由 $g \not\equiv -\infty$ 是凹函数, 同理可以得知 S_1 是凸集. 根据定理的条件, 不妨设 f 在某点 $x^0 \in \text{dom} f$ 处是连续的, 则由 $f(x^0) < +\infty$ 知, 存在 η_0 使

$$\eta_0 = f(x^0), \quad (2.2.19)$$

故由(2.2.18)有 $(x^0, \eta_0) \in S_2$. 因为 f 在点 x^0 处是连续的, 由定义 2.2.1 知存在 x^0 的邻域 $U(x^0)$, 对任意的 $\varepsilon > 0$, 有

$$|f(x) - f(x^0)| < \varepsilon \quad \forall x \in U(x^0).$$

注意到(2.2.19)得

$$\eta_0 - \varepsilon < f(x) < \eta_0 + \varepsilon \quad \forall x \in U(x^0),$$

于是存在某点 $(x', \eta') \in U(x^0) \times (\eta_0 - \varepsilon, \eta_0 + \varepsilon)$. 据此, 我们有 $(x', \eta') \in \text{int} S_2$, 从而得 $\text{int} S_2 \neq \emptyset$. 另外, 由 $x^0 \in \text{dom} g$ 知 $g(x^0) > -\infty$, 因而存在 $\eta'' = g(x^0) + m$. 由(2.2.17)知 $(x^0, \eta'') \in S_1$, 故得 $S_1 \neq \emptyset$. 现在再证明 $S_1 \cap \text{int} S_2 = \emptyset$. 事实上若不然, 则存在

$(\bar{x}, \bar{\eta}) \in S_1 \cap \text{int}S_2$, 于是 $(\bar{x}, \bar{\eta}) \in S_1$, 从而由(2.2.17)知

$$\bar{\eta} \leq g(\bar{x}) + m. \quad (2.2.20)$$

又由 $(\bar{x}, \bar{\eta}) \in \text{int}S_2$ 知存在 \bar{x} 的邻域 $U(\bar{x})$ 和 $\delta > 0$, 使 $U(\bar{x}) \times (\bar{\eta} - \delta, \bar{\eta} + \delta) \subset S_2$. 由此, 按(2.2.18)知 $f(\bar{x}) \leq \bar{\eta} - \delta$, 再由(2.2.20)得

$$f(\bar{x}) \leq \bar{\eta} - \delta < \bar{\eta} \leq g(\bar{x}) + m.$$

这说明存在 $\bar{x} \in \mathcal{X}$ 有 $m > f(\bar{x}) - g(\bar{x})$, 导致与(2.2.16)相矛盾. 从以上得到的 $S_1 \neq \emptyset$, $\text{int}S_2 \neq \emptyset$ 和 $S_1 \cap \text{int}S_2 \neq \emptyset$, 根据定理 1.3.7 的(1), 得知存在闭超平面正常分离 S_1 和 S_2 , 按定义 1.3.2 的(1)也即存在 $(x^*, \eta^*) \in (\mathcal{X}^* \times R) \setminus \{\theta\}$ 和 $\beta \in R$ 有

$$\langle (x^*, \eta^*), (x^1, \eta_1) \rangle \leq \beta \leq \langle (x^*, \eta^*), (x^2, \eta_2) \rangle$$

$$\forall (x^1, \eta_1) \in S_1, \forall (x^2, \eta_2) \in S_2,$$

即

$$\langle x^*, x^1 \rangle + \langle \eta^*, \eta_1 \rangle \leq \beta \leq \langle x^*, x^2 \rangle + \langle \eta^*, \eta_2 \rangle$$

$$\forall (x^1, \eta_1) \in S_1, \forall (x^2, \eta_2) \in S_2.$$

由(2.2.17)和(2.2.18)可取 $x = x^1 = x^2 \in \mathcal{X}$ 使 $\eta = f(x) = g(x) + m$, 则得

$$\langle x^*, x \rangle + \eta^* [g(x) + m] \leq \beta \leq \langle x^*, x \rangle + \eta^* f(x)$$

$$\forall x \in \mathcal{X}.$$

因为 $\eta^* \neq 0$, 记 $r = -\frac{\beta}{\eta^*} \in R$, $y^* = -\frac{x^*}{\eta^*} \in \mathcal{X}^*$, 从上式推知

$$\langle y^*, x \rangle - f(x) \leq r, \quad m + r \leq \langle y^*, x \rangle - g(x) \quad \forall x \in \mathcal{X},$$

于是由定义 2.2.4 的(1)和定义 2.2.5, 有

$$f^*(y^*) = \sup_{x \in \mathcal{X}} [\langle y^*, x \rangle - f(x)] \leq r,$$

$$m + r \leqslant g^*(y^*) = \inf_{x \in \mathcal{X}} [\langle y^*, x \rangle - g(x)].$$

据此, 我们得知 $m \leqslant g^*(y^*) - f^*(x^*)$, 再由 (2.2.15) 和 (2.2.16) 有

$$m \leqslant g^*(y^*) - f^*(x^*) \leqslant \sup_{x^* \in \mathcal{X}^*} \{g^*(x^*) - f^*(x^*)\} \leqslant m,$$

因而

$$m = \sup_{x^* \in \mathcal{X}^*} \{g^*(x^*) - f^*(x^*)\}.$$

最后, 注意到 (2.2.16), 定理得证. \square

考虑线性拓扑空间 \mathcal{X} 中凸集 S 上的指示函数

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathcal{X} \setminus S. \end{cases}$$

在例 2.1.1 中, 我们已指出它是凸函数. 从定义 2.2.4 的 (1) 可知, 有

$$\delta_S^*(x^*) = \sup_{x \in \mathcal{X}} \{\langle x^*, x \rangle - \delta_S(x)\} = \sup_{x \in S} \langle x^*, x \rangle,$$

即

$$\delta_S^*(x^*) = \sup_{x \in S} \langle x^*, x \rangle, \quad x^* \in \mathcal{X}^*. \quad (2.2.21)$$

据此, 由定义 1.3.4 得知 δ_S 的共轭函数 $\delta_S^*: \mathcal{X}^* \rightarrow R \cup \{+\infty\}$ 是 S 的支撑函数. 由定理 2.2.10 的 (1) 可知, δ_S^* 总是下半连续的正常凸函数.

定理 2.2.16 设 $S \subset \mathcal{X}$ 是闭凸集, $\text{int} S \neq \emptyset$, 则

$$S = \{x \in \mathcal{X} \mid \langle x^*, x \rangle \leqslant \delta_S^*(x^*) \quad \forall x^* \in \mathcal{X}^*\}. \quad (2.2.22)$$

证明 因为 δ_S^* 是 S 的支撑函数. 由定理 1.3.13, 注意到 S 是闭凸的即可得证. \square

定理 2.2.16 表明任一闭凸集可以由其上的指示函数的对偶函数来表示.

§ 2.3 广义凸函数

这一节,阐述广义意义下的凸函数及其有关性质.我们主要介绍两类重要的广义凸函数——拟凸函数类和伪凸函数类.

先介绍线性空间上的拟凸函数类.

设 \mathcal{V} 是线性空间.

定义 2.3.1 设 $S \subset \mathcal{V}$ 是非空凸集, $f: S \rightarrow R$ 是实值函数.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \max\{f(x^1), f(x^2)\} \quad \forall x^1, x^2 \in S, \quad (2.3.1)$$

则称 f 是集合 S 上的拟凸函数,或函数 f 在 S 上是拟凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} f(\lambda x^1 + (1 - \lambda)x^2) &< \max\{f(x^1), f(x^2)\} \\ \forall x^1, x^2 \in S, f(x^1) &\neq f(x^2), \end{aligned} \quad (2.3.2)$$

则称 f 是集合 S 上的严格拟凸函数,或函数 f 在 S 上是严格拟凸的.

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} f(\lambda x^1 + (1 - \lambda)x^2) &< \max\{f(x^1), f(x^2)\} \\ \forall x^1, x^2 \in S, x^1 &\neq x^2, \end{aligned} \quad (2.3.3)$$

则称 f 是集合 S 上的强拟凸函数,或函数 f 在 S 上是强拟凸的.

若 $-f$ 是 S 上的(严格,强)拟凸函数,则称 f 是 S 上的(严格,强)拟凹函数,或 f 在 S 上是(严格,强)拟凹的.

图 2.3.1 的(a)、(b)和(c)依次是单变量的拟凸函数、严格拟凸函数和强拟凸函数的图象.

注 2.3.1 在定义 2.3.1 中固定 $x^1 = x^0 \in S$, 令 $x^2 = x$, 则 (2.3.1) 成为

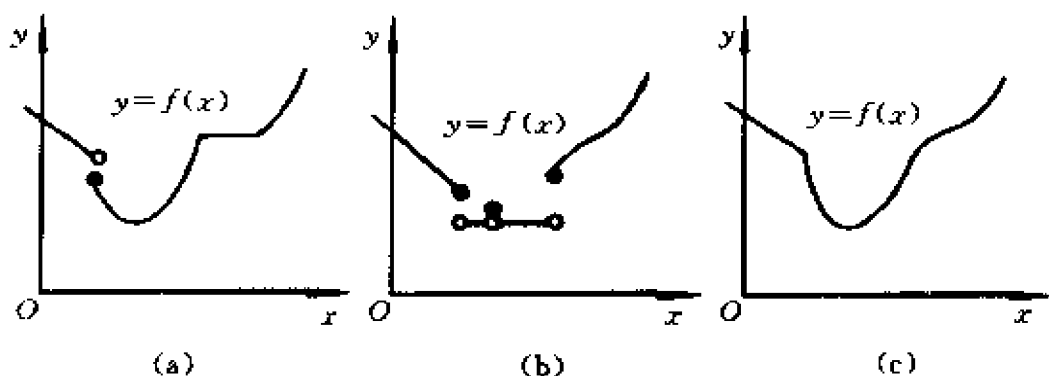


图 2.3.1

$$f(\lambda x^0 + (1 - \lambda)x) \leq \max\{f(x^0), f(x)\} \quad \forall x \in S,$$

这时,称 f 在点 x^0 处相对于集合 S 是拟凸的. 同样地,若 $x^1 = x^0$ 和 $x^2 = x$ 分别满足相应的(2.3.2)和(2.3.3),则称 f 在点 x^0 处相对于 S 是严格拟凸的和强拟凸的.

注 2.3.2 由定义 2.1.1 和定义 2.3.1 直接可以得知凸集上的凸函数是拟凸和严格拟凸的,严格凸函数是严格拟凸的,强凸函数是强拟凸的,反之,则不然. 从定义 2.3.1 还容易推知,强拟凸函数也是拟凸的和严格拟凸的. 然而,严格拟凸函数却不一定是拟凸的.

以下给出有关拟凸函数的两个基本结果.

定理 2.3.1 设 $S \subset \mathcal{V}$ 是非空凸集,则 $f: S \rightarrow R$ 是拟凸函数当且仅当对任何正整数 $m \geq 2$ 和任意的 $\lambda_i \geq 0$ ($i = 1, \dots, m$),

$$\sum_{i=1}^m \lambda_i = 1, \text{ 有}$$

$$f\left(\sum_{i=1}^m \lambda_i x^i\right) \leq \max\{f(x^1), \dots, f(x^m)\} \quad \forall x^1, \dots, x^m \in S.$$

证明 与定理 2.1.1 的证明类似,利用定义 2.3.1 的(1)可以推证. \square

定理 2.3.2 设 $S \subset \mathcal{V}$ 是非空凸集,则 $f: S \rightarrow R$ 是拟凸函数当且仅当对任意的 $c \in R$, 水平集 $H_S(f, c) = \{x \in S | f(x) \leq c\}$ 是

凸集.

证明 必要性. 任取 $x^1, x^2 \in H_S(f, c) \subset S$, 则有 $f(x^1) \leq c$ 和 $f(x^2) \leq c$. 因为 S 是凸集, 故对任意的 $\lambda \in (0, 1)$ 有 $\lambda x^1 + (1 - \lambda)x^2 \in S$. 由于 f 是 S 上的拟凸函数, 按定义 2.3.1 的(1)得

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \max\{f(x^1), f(x^2)\} \leq c,$$

于是 $\lambda x^1 + (1 - \lambda)x^2 \in H_S(f, c)$. 据此, 由定义 1.1.1 可知 $H_S(f, c)$ 是凸集.

充分性. 对任意的 $x^1, x^2 \in S$, 取

$$c = \max\{f(x^1), f(x^2)\}, \quad (2.3.4)$$

则 $x^1, x^2 \in H_S(f, c)$. 因为 $H_S(f, c)$ 是凸集, 故对任意的 $\lambda \in (0, 1)$, 有 $\lambda x^1 + (1 - \lambda)x^2 \in H_S(f, c)$, 于是得 $f(\lambda x^1 + (1 - \lambda)x^2) \leq c$. 据此, 由(2.3.4)按定义 2.3.1 的(1), 便知 f 在 S 上是拟凸的. \square

注 2.3.3 由定理 2.1.2 已知, 函数的相应任意水平集为凸集是该函数为凸函数的必要条件. 定理 2.3.2 则进一步表明了函数的相应任意水平集为凸集是该函数为拟凸的充要条件. 从而可见, 函数的拟凸性更充分地反映了其相应水平集的凸性特征.

下面的定理指出, 在线性拓扑空间上为下半连续的严格拟凸函数必定是拟凸的.

设 \mathcal{X} 是线性拓扑空间.

定理 2.3.3 设 $S \subset \mathcal{X}$ 是非空凸集. 若 $f: S \rightarrow R$ 是下半连续的严格拟凸函数, 则 f 是 S 上的拟凸函数.

证明 设对任意的 $x^1, x^2 \in S$, 有 $f(x^1) \neq f(x^2)$. 由 f 在 S 上是严格拟凸的, 从(2.3.2)可得(2.3.1). 因此, 根据定义 2.3.1 中的(1)即知 f 在 S 上是拟凸的.

设存在 $\bar{x}^1, \bar{x}^2 \in S$ 有 $f(\bar{x}^1) = f(\bar{x}^2)$. 我们证明这时条件(2.3.1)成立, 即对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda \bar{x}^1 + (1 - \lambda)\bar{x}^2) \leq f(\bar{x}^1) = f(\bar{x}^2).$$

为此,与定理 2.1.3 的证明类似,用反证法,假设对某个 $\lambda_0 \in (0, 1)$ 有

$$f(\lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2) > f(\bar{x}^1) = f(\bar{x}^2).$$

因为 S 是凸集,故 $x^0 = \lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2 \in S$, 由上式得

$$f(\bar{x}^1) < f(x^0), \quad f(\bar{x}^2) < f(x^0), \quad (2.3.5)$$

取充分小的 $\epsilon_0 > 0$, 使 (2.3.5) 的第一式有 $f(\bar{x}^1) + \epsilon_0 < f(x^0)$, 即

$$f(\bar{x}^1) < f(x^0) - \epsilon_0.$$

因为 f 在 $x^0 \in S$ 处是下半连续的, 由定义 2.1.2 的 (1) 知, 对任意的 $\epsilon > 0$, 存在 λ_0 的邻域 $U(\lambda_0)$, 有

$$-\epsilon < f(\lambda \bar{x}^1 + (1 - \lambda) \bar{x}^2) - f(x^0) \quad \forall \lambda \in U(\lambda_0).$$

据此, 由以上两式得知

$$f(\bar{x}^1) < f(x^0) - \epsilon_0 < f(\lambda \bar{x}^1 + (1 - \lambda) \bar{x}^2) \quad \forall \lambda \in U(\lambda_0).$$

取 $\lambda' \in U(\lambda_0)$, $\lambda' \neq \lambda_0$, 记 $x' = \lambda' \bar{x}^1 + (1 - \lambda') \bar{x}^2$, 则有

$$f(\bar{x}^1) < f(x'). \quad (2.3.6)$$

不妨设 $x^0 \in (\bar{x}^1, x')$, 则 $x' \in (x^0, \bar{x}^2)$. 因为 f 是严格拟凸的, 所以由定义 2.3.1 的 (2) 分别注意到 (2.3.6) 和 (2.3.5) 的第 2 式得到

$$f(x^0) < \max\{f(\bar{x}^1), f(x')\} = f(x'),$$

$$f(x') < \max\{f(x^0), f(\bar{x}^2)\} = f(x^0),$$

但这两式是矛盾的, 定理得证. \square

拟凸函数的以下一些性质与凸函数的相应性质相类似.

定理 2.3.4 设 $S \subset \mathcal{X}$ 是非空凸集, $f_i: S \rightarrow R$ 是拟凸函数 ($i \in I$, I 是指标集). 若

$$f(x) = \sup_{i \in I} f_i(x), \quad (2.3.7)$$

则 $f: S \rightarrow R$ 是拟凸函数.

证明 由 $f_i (i \in I)$ 是拟凸的, 根据定理 2.3.2 知, 对任意的 $c \in R$, 水平集 $H_S(f_i, c) = \{x \in S | f_i(x) \leq c\} (i \in I)$ 是凸的. 因此, 由定理 1.1.2 的 (2) 得知 $\bigcap_{i \in I} H_S(f_i, c)$ 是凸集. 任取 $x \in \bigcap_{i \in I} H_S(f_i, c)$, 则 $f_i(x) \leq c (i \in I)$, 由 (2.3.7), 它意味着 $f(x) \leq c$, 故得

$$H_S(f, c) = \{x \in S | f(x) \leq c\} = \bigcap_{i \in I} H_S(f_i, c).$$

由于对任意的 $c \in R$ 水平集 $H_S(f, c)$ 是凸的, 再根据定理 2.3.2 即得 f 是 S 上的拟凸函数. \square

定理 2.3.5 设 $f: \mathcal{V} \rightarrow R$ 和 $g: R \rightarrow R$ 是实值函数, 并且 $g(t) = f(x + td)$ (其中 $d \in \mathcal{V}$), 则 f 是拟凸函数当且仅当对任意的 $d \in \mathcal{V}$, g 是拟凸函数.

证明 利用定义 2.3.1 的 (1), 与定理 2.1.8 的证明类似可得证. \square

定理 2.3.6 设 $h: \mathcal{V} \rightarrow R$ 是拟凸函数, $g: R \rightarrow R$ 是非减函数, $f(x) = g(h(x))$, 则 $f: \mathcal{V} \rightarrow R$ 是拟凸函数.

证明 任取 $x^1, x^2 \in \mathcal{V}$ 和 $\lambda \in (0, 1)$, 因为 h 是拟凸的和 g 是非减的, 有

$$\begin{aligned} f(\lambda x^1 + (1 - \lambda)x^2) &= g(h(\lambda x^1 + (1 - \lambda)x^2)) \\ &\leq g(\max\{h(x^1), h(x^2)\}) \\ &= \max\{g(h(x^1)), g(h(x^2))\} \\ &= \max\{f(x^1), f(x^2)\}. \end{aligned}$$

由此, 根据定义 2.3.1 的 (1) 得知 f 是拟凸的. \square

注 2.3.4 定理 2.3.6 与定理 2.1.9 相似. 但要注意, 在定理 2.3.6 中并不要求 g 具有凸性或拟凸性.

再引进另一类由广义导数定义的广义凸函数的概念. 为此, 先给出映射的 Gâteaux 导数的定义.

再设 \mathscr{Y} 是线性拓扑空间.

定义 2.3.2 设 $\varphi: \mathscr{X} \rightarrow \mathscr{Y}$ 是映射, 点 $x^0 \in \mathscr{X}$. 若对任意的 $d \in \mathscr{X}$, 极限

$$\varphi'_{x^0}(d) = \lim_{t \rightarrow 0} \frac{\varphi(x^0 + td) - \varphi(x^0)}{t} \quad (2.3.8)$$

存在, 则称 φ 在点 x^0 处是 Gâteaux 可微的. 映射 $\varphi'_{x^0}: \mathscr{X} \rightarrow \mathscr{Y}$, $d \mapsto \varphi'_{x^0}(d)$ 称为是 φ 在点 x^0 处的 Gâteaux 导数. 若 φ 在集合 $S \subset \mathscr{X}$ 的每一点处都是 Gâteaux 可微的, 则称 φ 在 S 上是 Gâteaux 可微的.

特别地, 若 $\varphi: \mathscr{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 点 $x^0 \in \text{int}(\text{dom} \varphi)$, 则 φ 在点 x^0 处的 Gâteaux 导数 $\varphi'_{x^0}: \mathscr{X} \rightarrow R$ 是实值函数.

注 2.3.5 由定义 2.3.2 易知, 对于 $x \in \mathscr{X}$, 有

$$\varphi'_x(0) = 0, \quad (2.3.9)$$

$$\varphi'_x(\alpha d) = \alpha \varphi'_x(d), \alpha \in R. \quad (2.3.10)$$

定理 2.3.7 设 $f: \mathscr{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 点 $x^0 \in \text{int}(\text{dom} f)$. 若 f 在点 x^0 处是 Gâteaux 可微的, 则 f 在点 x^0 处是连续的.

证明 由 f 在点 x^0 处是 Gâteaux 可微的按定义 2.3.2 可知, 对任意的 $d \in \mathscr{X}$, 极限

$$f'_{x^0}(d) = \lim_{t \rightarrow 0} \frac{f(x^0 + td) - f(x^0)}{t}$$

存在. 据此, 我们有

$$f(x^0 + td) - f(x^0) = f'_{x^0}(d)t + o(t),$$

其中 $o(t) \rightarrow 0$ ($t \rightarrow 0$). 记 $x = x^0 + td$, 则对任意的 $\varepsilon > 0$, 只要 t 充分小便存在 x^0 的邻域 $U(x^0)$, 使

$$|f(x) - f(x^0)| = |f'_{x^0}(d)t + o(t)| < \varepsilon \quad \forall x \in U(x^0).$$

于是,由定义 2.2.1,即知 f 在点 x^0 处是连续的. \square

以下介绍线性拓扑空间上的伪凸函数类.

定义 2.3.3 设 $S \subset \mathscr{E}$ 是非空开凸集,实值函数 $f: S \rightarrow R$ 是 Gâteaux 可微的.

(1) 若有

$$f(x^1) > f(x^2) \Rightarrow f'_{x^1}(x^2 - x^1) < 0 \quad \forall x^1, x^2 \in S, \quad (2.3.11)$$

则称 f 是集合 S 上的伪凸函数,或函数 f 在 S 上是伪凸的.

(2) 若有

$$f(x^1) \geq f(x^2) \Rightarrow f'_{x^1}(x^2 - x^1) < 0 \quad \forall x^1, x^2 \in S, x^1 \neq x^2, \quad (2.3.12)$$

则称 f 是集合 S 上的强伪凸函数,或函数 f 在 S 上是强伪凸的.

(3) 若有

$$f(x^1) \geq f(x^2) \Rightarrow f'_{x^1}(x^2 - x^1) \leq 0 \quad \forall x^1, x^2 \in S, \quad (2.3.13)$$

则称 f 是集合 S 上的次伪凸函数,或函数 f 在 S 上是次伪凸的.

若 $-f$ 是 S 上的(强,次)伪凸函数,则称 f 是 S 上的(强,次)伪凹函数,或 f 在 S 上是(强,次)伪凹的.

图 2.3.2 的(a)、(b)和(c)依次是单变量的伪凸函数、强伪凸

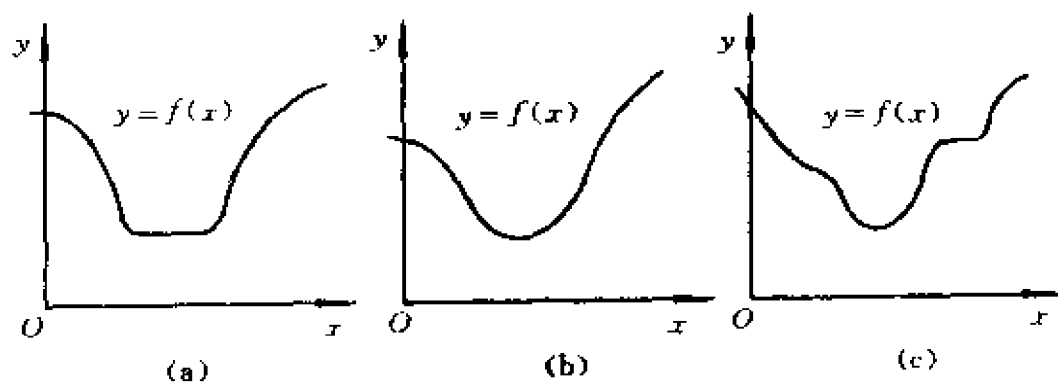


图 2.3.2

函数和次伪凸函数的图象.

注 2.3.6 在定义 2.3.3 中固定 $x^1 = x^0 \in S$, 令 $x^1 = x$, 则对应于定义 2.3.3 的(1)、(2)和(3), 依次称 f 在点 x^0 处相对于集合 S 是伪凸的、强伪凸的和次伪凸的.

注 2.3.7 由定义 2.3.3 不难得知, 若 f 是 S 上的强伪凸(凹)函数, 则 f 是 S 上的伪凸(凹)函数和次伪凸(凹)函数, 但反之不然.

注 2.3.8 对于有限维 Euclid 空间 R^n 中开凸集 S 上的可微函数 $f: S \rightarrow R$, 由定义 2.3.2 可知, f 在点 x^1 处的 Gâteaux 导数在 $d = x^2 - x^1$ 处的值 $f'_{x^1}(x^2 - x^1)$, 即为 f 在点 x^1 处沿方向 $d = x^2 - x^1$ 的方向导数, 故有 $f'_{x^1}(x^2 - x^1) = \nabla f(x^1)^T(x^2 - x^1)$ (其中 $\nabla f(x^1)$ 是 f 在点 x^1 处的梯度). 这时, f 在 S 上是伪凸的意味着

$$f(x^1) > f(x^2) \Rightarrow \nabla f(x^1)^T(x^2 - x^1) < 0 \quad \forall x^1, x^2 \in S,$$

f 在 S 上是强伪凸的意味着

$$f(x^1) \geq f(x^2) \Rightarrow \nabla f(x^1)^T(x^2 - x^1) < 0$$

$$\forall x^1, x^2 \in S, x^1 \neq x^2,$$

f 在 S 上是次伪凸的意即

$$f(x^1) \geq f(x^2) \Rightarrow \nabla f(x^1)^T(x^2 - x^1) \leq 0 \quad \forall x^1, x^2 \in S.$$

下面的定理指出, 线性拓扑空间上 Gâteaux 可微的凸函数和强凸函数分别也是伪凸函数和强伪凸函数.

定理 2.3.8 设 $S \subset \mathcal{X}$ 是非空开凸集, $f: S \rightarrow R$ 是 Gâteaux 可微的.

(1) 若 f 是 S 上的凸函数, 则 f 是 S 上的伪凸函数和次伪凸函数.

(2) 若 f 是 S 上的强凸函数, 则 f 是 S 上的强伪凸函数.

证明 (1) 因 f 是 S 上的凸函数, 按定义 2.1.1 的(1), 对于

任意的 $x^1, x^2 \in S$ 和任意的 $t \in (0, 1)$, 有

$$\begin{aligned} f(x^1 + t(x^2 - x^1)) - f(x^1) &= f(tx^2 + (1-t)x^1) - f(x^1) \\ &\leq tf(x^2) + (1-t)f(x^1) - f(x^1) \\ &= t[f(x^2) - f(x^1)]. \end{aligned}$$

由此,按定义 2.3.2 得

$$\begin{aligned} f'_{x^1}(x^2 - x^1) &= \lim_{t \rightarrow 0} \frac{f(x^1 + t(x^2 - x^1)) - f(x^1)}{t} \\ &\leq f(x^2) - f(x^1). \end{aligned} \quad (2.3.14)$$

依据(2.3.14),从 $f(x^1) > f(x^2)$ 即可推出 $f'_{x^1}(x^2 - x^1) < 0$, 于是按定义 2.3.3 的(1)得知, f 是 S 上的伪凸函数.

又据(2.3.14),从 $f(x^1) \geq f(x^2)$ 可以推出 $f'_{x^1}(x^2 - x^1) \leq 0$, 故按定义 2.3.3 的(3),即知 f 是 S 上的次伪凸函数.

(2) 与(1)的证明类似,由定义 2.3.2 和定义 2.1.1 的(3)可以推知,对任意的 $x^1, x^2 \in S, x^1 \neq x^2$, 有

$$f'_{x^1}(x^2 - x^1) < f(x^2) - f(x^1).$$

由此,从 $f(x^1) \geq f(x^2)$ 可推出 $f'_{x^1}(x^2 - x^1) < 0$, 故由定义 2.3.3 中的(2),即得 f 是 S 上的强伪凸函数. \square

以下讨论函数的伪凸性与拟凸性之间的几个关系.

定理 2.3.9 设 $S \subset \mathscr{R}^n$ 是非空开凸集, $f: S \rightarrow R$ 是 Gâteaux 可微的.

(1) 若 f 是 S 上的伪凸函数,则 f 是 S 上的严格拟凸函数.

(2) 若 f 是 S 上的伪凸函数,则 f 是 S 上的拟凸函数.

证明 (1) 用反证法,假设 f 在 S 上不是严格拟凸的,则由定义 2.3.1 的(2)得知,存在 $x^1, x^2 \in S, f(x^1) \neq f(x^2)$, 对某 $\lambda_0 \in (0, 1)$ 有

$$f(\lambda_0 x^1 + (1 - \lambda_0)x^2) \geq \max\{f(x^1), f(x^2)\}. \quad (2.3.15)$$

不妨设

$$f(\mathbf{x}^1) > f(\mathbf{x}^2), \quad (2.3.16)$$

则从(2.3.15)知

$$f(\lambda_0 \mathbf{x}^1 + (1 - \lambda_0) \mathbf{x}^2) \geq f(\mathbf{x}^1) > f(\mathbf{x}^2). \quad (2.3.17)$$

由 S 是凸集, f 在闭线段 $[\mathbf{x}^1, \mathbf{x}^2] \subset S$ 上是 Gâteaux 可微的, 根据定理 2.3.7, 它在该线段上是连续的. 因此, f 在 $[\mathbf{x}^1, \mathbf{x}^2]$ 上有最大值. 从(2.3.17)和 $\lambda_0 \mathbf{x}^1 + (1 - \lambda_0) \mathbf{x}^2 \in (\mathbf{x}^1, \mathbf{x}^2)$ 可知, 该最大值可在 $(\mathbf{x}^1, \mathbf{x}^2)$ 内取到, 于是存在点 $\bar{\mathbf{x}} \in (\mathbf{x}^1, \mathbf{x}^2)$, 对任意的 $t \in (0, 1)$ 有

$$f(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}} + t(\mathbf{x}^1 - \bar{\mathbf{x}})),$$

$$f(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}} + t(\mathbf{x}^2 - \bar{\mathbf{x}})).$$

据此, 由(2.3.8)得

$$f'_{\bar{\mathbf{x}}}(\mathbf{x}^1 - \bar{\mathbf{x}}) = \lim_{t \rightarrow 0} \frac{f(\bar{\mathbf{x}} + t(\mathbf{x}^1 - \bar{\mathbf{x}})) - f(\bar{\mathbf{x}})}{t} \leq 0, \quad (2.3.18)$$

$$f'_{\bar{\mathbf{x}}}(\mathbf{x}^2 - \bar{\mathbf{x}}) = \lim_{t \rightarrow 0} \frac{f(\bar{\mathbf{x}} + t(\mathbf{x}^2 - \bar{\mathbf{x}})) - f(\bar{\mathbf{x}})}{t} \leq 0. \quad (2.3.19)$$

设 $\bar{\mathbf{x}} = \bar{\lambda} \mathbf{x}^1 + (1 - \bar{\lambda}) \mathbf{x}^2$ (其中 $\bar{\lambda} \in (0, 1)$), 则

$$\mathbf{x}^1 - \bar{\mathbf{x}} = -\frac{1 - \bar{\lambda}}{\bar{\lambda}} (\mathbf{x}^2 - \bar{\mathbf{x}}).$$

于是, 由(2.3.18)和(2.3.19)利用(2.3.10)有

$$\begin{aligned} 0 &\geq f'_{\bar{\mathbf{x}}}(\mathbf{x}^1 - \bar{\mathbf{x}}) = f'_{\bar{\mathbf{x}}}\left(-\frac{1 - \bar{\lambda}}{\bar{\lambda}}(\mathbf{x}^2 - \bar{\mathbf{x}})\right) \\ &= -\frac{1 - \bar{\lambda}}{\bar{\lambda}} f'_{\bar{\mathbf{x}}}(\mathbf{x}^2 - \bar{\mathbf{x}}) \geq 0, \end{aligned}$$

从而得

$$f'_{\bar{x}}(\mathbf{x}^2 - \bar{\mathbf{x}}) = 0. \quad (2.3.20)$$

因为已知 f 在 S 上是伪凸的, 所以由 (2.3.20) 可推知 $f(\bar{\mathbf{x}}) \leq f(\mathbf{x}^2)$ (否则, 若 $f(\bar{\mathbf{x}}) > f(\mathbf{x}^2)$, 由定义 2.3.3 的 (1) 得 $f'_{\bar{x}}(\mathbf{x}^2 - \bar{\mathbf{x}}) < 0$, 与 (2.3.20) 矛盾). 注意到 (2.3.16) 得 $f(\mathbf{x}^1) > f(\bar{\mathbf{x}})$, 因为 f 是连续的, 这与 $f(\bar{\mathbf{x}})$ 是 $(\mathbf{x}^1, \mathbf{x}^2)$ 中的最大值相矛盾.

(2) 由 f 在 S 上是 Gâteaux 可微的, 它在 S 上是连续的, 从而也是下半连续的. 于是, 由已证得的 (1) 和定理 2.3.3, 即知 f 在 S 上是拟凸的. \square

定理 2.3.10 设 $S \subset \mathcal{R}$ 是非空开凸集, $f: S \rightarrow R$ 是 Gâteaux 可微的. 若 f 是 S 上的强伪凸函数, 则 f 是 S 上的强拟凸函数.

证明 用反证法, 假设 f 在 S 上不是强拟凸的, 则由定义 2.3.1 的 (3) 得知, 存在 $\bar{\mathbf{x}}^1, \bar{\mathbf{x}}^2 \in S, \bar{\mathbf{x}}^1 \neq \bar{\mathbf{x}}^2$, 对某个 $\bar{\lambda} \in (0, 1)$ 有

$$f(\bar{\lambda} \bar{\mathbf{x}}^1 + (1 - \bar{\lambda}) \bar{\mathbf{x}}^2) \geq \max \{f(\bar{\mathbf{x}}^1), f(\bar{\mathbf{x}}^2)\}. \quad (2.3.21)$$

设 $\bar{\mathbf{x}} = \bar{\lambda} \bar{\mathbf{x}}^1 + (1 - \bar{\lambda}) \bar{\mathbf{x}}^2 \neq \bar{\mathbf{x}}^1$, 由 (2.3.21) 有 $f(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}}^1)$, 于是根据 f 是强伪凸函数, 按定义 2.3.3 的 (2) 得

$$f'_{\bar{x}}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}) < 0. \quad (2.3.22)$$

同理, 从 (2.3.21) 有 $f(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}}^2)$, 从而又得

$$f'_{\bar{x}}(\bar{\mathbf{x}}^2 - \bar{\mathbf{x}}) < 0. \quad (2.3.23)$$

因为 $\bar{\mathbf{x}}^1 - \bar{\mathbf{x}} = -\frac{1 - \bar{\lambda}}{\bar{\lambda}}(\bar{\mathbf{x}}^2 - \bar{\mathbf{x}})$, 所以由 (2.3.22) 和 (2.3.23) 利用 (2.3.10) 推知

$$0 > f'_{\bar{x}}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}) = -\frac{1 - \bar{\lambda}}{\bar{\lambda}} f'_{\bar{x}}(\bar{\mathbf{x}}^2 - \bar{\mathbf{x}}) > 0,$$

导致矛盾. \square

定理 2.3.11 设 $S \subset \mathcal{R}$ 是非空开凸集, $f: S \rightarrow R$ 是

Gâteaux 可微的, 则 f 是 S 上的拟凸函数当且仅当 f 是 S 上的次伪凸函数.

证明 必要性. 由 f 在 S 上是拟凸的, 按定义 2.3.1 的(1)知, 对任意的 $\mathbf{x}^1, \mathbf{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda \mathbf{x}^2 + (1 - \lambda)\mathbf{x}^1) \leq \max\{f(\mathbf{x}^2), f(\mathbf{x}^1)\}.$$

设 $f(\mathbf{x}^1) \geq f(\mathbf{x}^2)$, 从上式知

$$\begin{aligned} f(\mathbf{x}^1 + \lambda(\mathbf{x}^2 - \mathbf{x}^1)) &= f(\mathbf{x}^1) \\ &= f(\lambda \mathbf{x}^2 + (1 - \lambda)\mathbf{x}^1) = f(\mathbf{x}^1) \leq 0, \end{aligned}$$

据此, 由定义 2.3.2 得到

$$f'_{\mathbf{x}^1}(\mathbf{x}^2 - \mathbf{x}^1) = \lim_{\lambda \rightarrow 0} \frac{f(\mathbf{x}^1 + \lambda(\mathbf{x}^2 - \mathbf{x}^1)) - f(\mathbf{x}^1)}{\lambda} \leq 0.$$

于是, 按定义 2.3.3 的(3)即知 f 在 S 上是次伪凸的.

充分性. 已知 f 在 S 上是次伪凸的, 用反证法, 假设 f 在 S 上不是拟凸的. 从后者由定义 2.3.1 中的(1)可知, 存在 $\bar{\mathbf{x}}^1, \bar{\mathbf{x}}^2 \in S$ 和 $\lambda_0 \in (0, 1)$ 有

$$f(\lambda_0 \bar{\mathbf{x}}^1 + (1 - \lambda_0) \bar{\mathbf{x}}^2) > \max\{f(\bar{\mathbf{x}}^1), f(\bar{\mathbf{x}}^2)\}. \quad (2.3.24)$$

因为 S 是凸集, 故

$$\mathbf{x}^0 = \lambda_0 \bar{\mathbf{x}}^1 + (1 - \lambda_0) \bar{\mathbf{x}}^2 \in S. \quad (2.3.25)$$

现设 $f(\bar{\mathbf{x}}^1) \geq f(\bar{\mathbf{x}}^2)$, 从(2.3.24)和(2.3.25)可得

$$f(\bar{\mathbf{x}}^2) \leq f(\bar{\mathbf{x}}^1) < f(\mathbf{x}^0). \quad (2.3.26)$$

由于 f 在 S 上是次伪凸的, 根据(2.3.26)和定义 2.3.3 的(3)推知

$$f'_{\mathbf{x}^0}(\bar{\mathbf{x}}^1 - \mathbf{x}^0) \leq 0, \quad f'_{\mathbf{x}^0}(\bar{\mathbf{x}}^2 - \mathbf{x}^0) \leq 0.$$

由 (2.3.25), 有 $\bar{x}^1 - x^0 = (1 - \lambda_0)(\bar{x}^1 - \bar{x}^2)$ 和 $\bar{x}^2 - x^0 = \lambda_0(\bar{x}^2 - \bar{x}^1)$, 代入以上两式, 注意到 (2.3.10) 有 $f'_{x^0}(\bar{x}^1 - \bar{x}^2) \leq 0$ 和 $f'_{x^0}(\bar{x}^2 - \bar{x}^1) \leq 0$, 从而得

$$f'_{x^0}(\bar{x}^2 - \bar{x}^1) = 0. \quad (2.3.27)$$

记

$$\lambda' = \inf \{ \lambda \in [0, 1] \mid f(\lambda \bar{x}^1 + (1 - \lambda)x^0) \leq f(\bar{x}^1) \}, \quad (2.3.28)$$

则有 $\lambda' > 0$. 事实上, 若有 $\lambda' = 0$, 则由 f 在 S 上是 Gâteaux 可微的, 根据定理 2.3.7, 它在 S 上是连续的, 从而也是下半连续的. 由定理 2.2.5 的 (1) 和 (3) 知, $D = \{x \in S \mid f(x) \leq f(\bar{x}^1)\}$ 是闭集. 据此, 设 $\{\lambda_k\} \subset [0, 1]$, $\lambda_k \bar{x}^1 + (1 - \lambda_k)x^0 \in D$, 并且 $\lambda_k \rightarrow \lambda' = 0$ ($k \rightarrow \infty$), 则 $\lambda_k \bar{x}^1 + (1 - \lambda_k)x^0 \rightarrow x^0 \in D$, 即有 $f(x^0) \leq f(\bar{x}^1)$, 它与 (2.3.26) 相矛盾. 现令

$$x' = \lambda' \bar{x}^1 + (1 - \lambda')x^0, \quad (2.3.29)$$

则 $x' \in (\bar{x}^1, x^0)$, 并且由 (2.3.28) 有

$$f(x') \leq f(\bar{x}^1). \quad (2.3.30)$$

另外, 由 (2.3.28) 知

$$f(x) > f(\bar{x}^1) \quad \forall x \in (x', x^0). \quad (2.3.31)$$

从 f 在 S 上是 Gâteaux 可微的, 根据中值定理^[6], 有

$$f(x^0) - f(x') = f'_x(x^0 - x'),$$

其中 $\bar{x} \in (x', x^0)$. 由 (2.3.25) 和 (2.3.29) 知 $x^0 - x' = (1 - \lambda_0)\lambda'(\bar{x}^2 - \bar{x}^1)$. 代入上式并注意 (2.3.10), 得

$$f(x^0) - f(x') = (1 - \lambda_0)\lambda' f'_x(\bar{x}^2 - \bar{x}^1). \quad (2.3.32)$$

从 $\bar{x} \in (x', x^0)$ 和 (2.3.31) 知 $f(\bar{x}^1) < f(\bar{x})$, 又 \bar{x} 也是 \bar{x}^1 和 \bar{x}^2 的凸组合, 因此对 (2.3.26) 至 (2.3.27) 以 \bar{x} 替换 x^0 , 可以推得

$$f'_x(\bar{x}^2 - \bar{x}^1) = 0.$$

将它代入(2.3.32),再注意(2.3.30),便得 $f(x^0) = f(x') \leq f(\bar{x}^1)$, 这导致与(2.3.26)矛盾. \square

§ 2.4 凸映射和广义凸映射

现在将函数的凸性概念推广到映射的情况,论述凸映射和广义凸映射及其有关性质和关系. 由于映射值空间上的序由该空间中的锥来确定,故论述映射的凸性要依赖于相应的序锥. 因此,有关凸映射也称为锥凸映射.

设 \mathcal{V} 是线性空间, \mathcal{U} 是序线性拓扑空间. 除特别说明外, $K \subset \mathcal{U}$ 是内部非空的尖闭凸锥, \mathcal{U} 中的序由 K 确定.

先介绍锥凸映射类.

定义 2.4.1 设 $S \subset \mathcal{V}$ 是非空凸集, $\varphi: S \rightarrow \mathcal{U}$ 是映射.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K \quad (2.4.1) \\ \forall x^1, x^2 \in S,$$

则称 φ 是集合 S 上的 K -凸映射, 或映射 φ 在 S 上是 K -凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \\ \in \text{int}K \setminus \{0\} \quad \forall x^1, x^2 \in S, \varphi(x^1) \neq \varphi(x^2), \quad (2.4.2)$$

则称 φ 是集合 S 上的 K -严格凸映射 (K -次严格凸映射), 或映射 φ 在 S 上是 K -严格凸的 (K -次严格凸的).

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \\ \in \text{int}K \quad \forall x^1, x^2 \in S, x^1 \neq x^2, \quad (2.4.3)$$

则称 φ 是集合 S 上的 K -强凸映射, 或映射 φ 在 S 上是 K -强凸的.

注 2.4.1 由定义 2.4.1 直接可知, 凸集上的 K -强凸映射也是 K -凸的, K -严格凸的和 K -次严格凸的, 反之不然. 但是, K -严格凸映射却不一定是 K -凸映射.

注 2.4.2 特别地, 在定义 2.4.1 中, 当 $\mathscr{U} = R^m$ 和 $K = R_+^m$ ($m \geq 1$) 时, $\varphi = f: S \rightarrow R^m$ 是向量函数. f 在 S 上是 R_+^m -凸的, 即对任意的 $\lambda \in (0, 1)$ 有

$$\lambda f(x^1) + (1 - \lambda)f(x^2) - f(\lambda x^1 + (1 - \lambda)x^2) \in R_+^m,$$

或即

$$f(\lambda x^1 + (1 - \lambda)x^2)$$

$$\leq \lambda f(x^1) + (1 - \lambda)f(x^2) \quad \forall x^1, x^2 \in S. \quad (2.4.4)$$

这时, 称 f 为 S 上的凸向量函数. 将上述向量不等式 (2.4.4) 与 (2.1.1) 对比可知, f 是 S 上的凸向量函数也即它的每一分量函数在 S 上是凸的. 同样地, $f: S \rightarrow R^m$ 在 S 上是 R_+^m -严格凸的 (R_+^m -强凸的), 即对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda x^1 + (1 - \lambda)x^2) < \lambda f(x^1) + (1 - \lambda)f(x^2)$$

$$\forall x^1, x^2 \in S, f(x^1) \neq f(x^2) (x^1 \neq x^2),$$

并称 f 是 S 上的严格 (强) 凸向量函数. 这时, f 的每一分量函数在 S 上是严格 (强) 凸的. 显然, $f: S \rightarrow R^m$ 在 S 上是 R_+^m -次严格凸的, 也即它的每一分量函数是凸的, 并且至少有一是严格凸的. 若 $\mathscr{U} = R$ 和 $K = R_+$, 则由定义 2.4.1 和定义 2.1.1 可知, $f: S \rightarrow R$ 是 R_+ -凸的, 即 f 是 S 上的凸函数, $f: S \rightarrow R$ 是 R_+ -严格凸的和 R_+ -次严格凸的, 即 f 是 S 上的严格凸函数, 而 $f: S \rightarrow R$ 是 R_+ -强凸的, 即 f 是 S 上的强凸函数.

以下给出有关 K -凸映射的几个基本关系.

定理 2.4.1 设 $S \subset \mathscr{X}$ 是非空凸集, $\varphi: S \rightarrow \mathscr{U}$ 是映射.

- (1) 若 φ 是 S 上的 K -凸映射, 则 $\varphi(S) \subset \mathscr{U}$ 是 K -凸集.
- (2) 若 φ 是 S 上的 K -强凸映射, 则 $\varphi(S) \subset \mathscr{U}$ 是严格 K -

凸集.

证明 (1) 由 $S \neq \emptyset$ 知 $\varphi(S) \neq \emptyset$. 若 $\varphi(S)$ 是单点集, 由定义 1.4.4 可知它是 K -凸集. 现设 $\varphi(S)$ 至少包含两个点, 任取 $y^1, y^2 \in \varphi(S)$, 则存在 $x^1, x^2 \in S$, 使

$$y^1 = \varphi(x^1), y^2 = \varphi(x^2). \quad (2.4.5)$$

因为 S 是凸集, 故对任意的 $\lambda \in (0, 1)$ 有

$$\lambda x^1 + (1 - \lambda)x^2 \in S. \quad (2.4.6)$$

又 φ 是 S 上的 K -凸映射, 由定义 2.4.1 的(1)知

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K.$$

据此, 由(2.4.5)和(2.4.6)得到

$$\lambda y^1 + (1 - \lambda)y^2 \in \varphi(S) + K,$$

所以按定义 1.4.4 便知 $\varphi(S)$ 是 K -凸的.

(2) 与(1)的证明类似. 利用定义 2.4.1 的(3)和定义 1.4.5 即可得证. \square

定理 2.4.2 设 $S \subset \mathcal{V}$ 是非空凸集, 则 $\varphi: S \rightarrow \mathcal{W}$ 是 K -凸映射当且仅当对任何正整数 $m \geq 2$ 和任意的 $\lambda_i \geq 0$ ($i = 1, \dots, m$),

$\sum_{i=1}^m \lambda_i = 1$, 有

$$\sum_{i=1}^m \lambda_i \varphi(x^i) - \varphi\left(\sum_{i=1}^m \lambda_i x^i\right) \in K \quad \forall x^1, \dots, x^m \in S. \quad (2.4.7)$$

证明 充分性. 在(2.4.7)中取 $m = 2$, 由定义 2.4.1 中的(1)即知 φ 是 S 上的 K -凸映射.

必要性. 用数学归纳法, 与定理 2.1.1 证明中的必要性部分类似, 由定义 2.4.1 的(1)注意到 K 是凸锥即可得证. \square

定理 2.4.3 设 $S \subset \mathcal{V}$ 是非空凸集, 若 $\varphi: S \rightarrow \mathcal{W}$ 是 K -凸映

射, 则对任意的 $c \in \mathscr{U}$, K -水平集 $H_S(\varphi, c)_K = \{x \in S | c - \varphi(x) \in K\}$ 是凸集.

证明 任取 $x^1, x^2 \in H_S(\varphi, c)_K \subset S$, 则有

$$c - \varphi(x^1), c - \varphi(x^2) \in K. \quad (2.4.8)$$

因为 S 是凸集, 故对任意的 $\lambda \in (0, 1)$, 有 $\lambda x^1 + (1 - \lambda)x^2 \in S$. 由于 φ 是 S 上的 K -凸映射, 按定义 2.4.1 的(1)知

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K.$$

据此, 由(2.4.8)注意到 K 是凸锥, 有

$$\begin{aligned} c - \varphi(\lambda x^1 + (1 - \lambda)x^2) &\in c + K - \lambda\varphi(x^1) - (1 - \lambda)\varphi(x^2) \\ &= K + \lambda[c - \varphi(x^1)] + (1 - \lambda)[c - \varphi(x^2)] \\ &\subset K + \lambda K + (1 - \lambda)K \\ &= K + K \subset K. \end{aligned}$$

于是, 得到 $\lambda x^1 + (1 - \lambda)x^2 \in H_S(\varphi, c)_K$, 因而由定义 1.1.1 即知 $H_S(\varphi, c)_K$ 是凸集. \square

注 2.4.3 当 $\mathscr{U} = R$ 和 $K = R_+$ 时, $\varphi = f: S \rightarrow R$ 是实值函数, R_+ -水平集即水平集:

$$\begin{aligned} H_S(f, c)_{R_+} &= \{x \in S | c - f(x) \in R_+\} \\ &= \{x \in S | f(x) \leq c\} = H_S(f, c). \end{aligned}$$

这时, 定理 2.4.2 和定理 2.4.3 分别即是定理 2.1.1 和定理 2.1.2.

为讨论 K -凸映射和 K -严格凸映射之间的关系, 下面引入映射的 K -半连续性概念. 设 \mathscr{X} 是线性拓扑空间.

定义 2.4.2 设集合 $S \subset \mathscr{X}$ 非空, $\varphi: S \rightarrow \mathscr{U}$ 是映射, $K \subset \mathscr{U}$ 是非平凡锥. 若对任意的 $c \in \mathscr{U}$, K -水平集

$$H_S(\varphi, c)_K = \{x \in S | c - \varphi(x) \in \text{cl}K\}$$

是闭集, 则称 φ 是 S 上的 K -半连续映射, 或映射 φ 在 S 上是 K -半连续的.

注 2.4.4 设 $\mathscr{U} = R$ 和 $K = R_+$, 由定义 2.4.2 可知 $\varphi = f$ 在 S 上是 R_+ -半连续的, 意味着 R_+ -水平集 $H_S(f, c)_{R_+} = \{x \in S \mid c - f(x) \in R_+\}$, 也即水平集 $H_S(f, c) = \{x \in S \mid f(x) \leq c\}$ 是闭的. 由定理 2.2.5 的(1)和(3), 这就是函数 f 在 S 上是下半连续的. 同理, f 在 S 上是 R_- -半连续的, 意即 $\{x \in S \mid f(x) \geq c\}$ 是闭集, 于是由定理 2.2.6 的(1)和(3), 也即函数 f 在 S 上是上半连续的.

定理 2.4.4 设 $S \subset \mathscr{X}$ 是非空凸集. 若 $\varphi: S \rightarrow \mathscr{U}$ 是 K -半连续的 K -严格凸映射, 则 φ 是 S 上的 K -凸映射.

证明 设对任意的 $x^1, x^2 \in S$ 有 $\varphi(x^1) \neq \varphi(x^2)$. 由 φ 在 S 上是 K -严格凸映射, 按定义 2.4.1 的(2)有

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in \text{int}K \subset K.$$

因此, 由定义 2.4.1 的(1), 即知 φ 是 K -凸映射.

设存在 $\bar{x}^1, \bar{x}^2 \in S$ 使 $\varphi(\bar{x}^1) = \varphi(\bar{x}^2)$. 用反证法, 假设 φ 在 S 上不是 K -凸映射, 则按定义 2.4.1 的(1)可知, 存在 $\lambda_0 \in (0, 1)$, 有

$$\lambda_0\varphi(\bar{x}^1) + (1 - \lambda_0)\varphi(\bar{x}^2) - \varphi(\lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2) \notin K.$$

令 $x^0 = \lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2$, 则得

$$\varphi(\bar{x}^1) - \varphi(x^0) \notin K. \quad (2.4.9)$$

记

$$\lambda' = \inf \{ \lambda \in (\lambda_0, 1] \mid \varphi(\bar{x}^1) - \varphi(\lambda \bar{x}^1 + (1 - \lambda) \bar{x}^2) \in K \}, \quad (2.4.10)$$

则有 $\lambda' > \lambda_0$ (否则, 若有 $\lambda' = \lambda_0$, 由 φ 在 S 上是 K -半连续的, 按定义 2.4.2 知 $H_S(\varphi, \varphi(\bar{x}^1))_K = \{x \in S \mid \varphi(\bar{x}^1) - \varphi(x) \in K\}$ 是闭集. 由此, 设 $\{\lambda_k\} \subset (\lambda_0, 1]$, $\lambda_k \bar{x}^1 + (1 - \lambda_k) \bar{x}^2 \in H_S(\varphi, \varphi(\bar{x}^1))_K$, $\lambda_k \rightarrow \lambda' = \lambda_0 (k \rightarrow \infty)$, 则有 $\lambda_k \bar{x}^1 + (1 - \lambda_k) \bar{x}^2 \rightarrow \lambda_0 \bar{x}^1 + (1 - \lambda_0) \bar{x}^2$

$\in H_s(\varphi, \varphi(\bar{x}^1))_K$, 故得 $\varphi(\bar{x}^1) - \varphi(x^0) \in K$, 这与 (2.4.9) 相矛盾). 现令

$$x' = \lambda' \bar{x}^1 + (1 - \lambda') \bar{x}^2,$$

由 (2.4.10) 和 $\lambda' > \lambda_0$ 知

$$\varphi(\bar{x}^1) - \varphi(x) \notin K \quad \forall x \in (x', x^0). \quad (2.4.11)$$

任取 $\bar{x} \in (x', x^0)$, 则 $\bar{x} \in (\bar{x}^1, x^0)$. 记 $\bar{x} = \mu \bar{x}^1 + (1 - \mu)x^0$ ($\mu \in (0, 1)$), 由 (2.4.9) 有 $\varphi(\bar{x}^1) \neq \varphi(x^0)$, 因为 φ 是 K -严格凸映射, 据定义 2.4.1 的 (2) 知

$$\varphi(\bar{x}) \in \mu\varphi(\bar{x}^1) + (1 - \mu)\varphi(x^0) - \text{int}K. \quad (2.4.12)$$

另外, 因为 $x^0 \in (\bar{x}, \bar{x}^2)$, 记 $x^0 = \nu \bar{x} + (1 - \nu)\bar{x}^2$ ($\nu \in (0, 1)$), 从 (2.4.11) 有 $\varphi(\bar{x}) \neq \varphi(\bar{x}^1) = \varphi(\bar{x}^2)$, 所以由 φ 是 K -严格凸的又知

$$\varphi(x^0) \in \nu\varphi(\bar{x}) + (1 - \nu)\varphi(\bar{x}^2) - \text{int}K,$$

即

$$\varphi(x^0) \in \nu\varphi(\bar{x}) + (1 - \nu)\varphi(\bar{x}^1) - \text{int}K. \quad (2.4.13)$$

作 $\nu(2.4.12) + (2.4.13)$, 有

$$\begin{aligned} \nu\varphi(\bar{x}) + \varphi(x^0) &\in [\nu\mu\varphi(\bar{x}^1) + \nu(1 - \mu)\varphi(x^0) - \text{int}K] \\ &\quad + [\nu\varphi(\bar{x}) + (1 - \nu)\varphi(\bar{x}^1) - \text{int}K], \end{aligned}$$

即

$$(1 - \nu + \nu\mu)\varphi(x^0) \in (1 - \nu + \nu\mu)\varphi(\bar{x}^1) - \text{int}K,$$

从而得 $\varphi(\bar{x}^1) - \varphi(x^0) \in \text{int}K$, 这导致与 (2.4.9) 相矛盾. \square

注 2.4.5 当 $\mathcal{U} = R$ 和 $K = R_+$ 时, 定理 2.4.4 即定理 2.1.3.

以下阐述锥拟凸映射类.

先引进向量组的锥界集概念, 借助它给出锥拟凸映射类的定义.

定义 2.4.3 设 $y^1, \dots, y^m \in \mathscr{Y} (m \geq 2)$. 若 $\bar{y} \in \bigcap_{i=1}^m (y^i + K)$, 并且不存在 $y \in \bigcap_{i=1}^m (y^i + K)$ 使得

$$\bar{y} - y \in K \setminus \{\theta\},$$

则称 \bar{y} 是向量组 $\{y^1, \dots, y^m\}$ 的 K -界点. $\{y^1, \dots, y^m\}$ 的所有 K -界点组成的集合称为它的 K -界集, 记作 $K\text{-bou}\{y^1, \dots, y^m\}$.

注 2.4.6 对于给定的 $y^1, \dots, y^m \in \mathscr{Y}$, 由 $\text{int}K \neq \emptyset$ 可以推得 $\bigcap_{i=1}^m (y^i + K) \neq \emptyset$, 并且由 K 是尖闭凸锥不难得知 $\bigcap_{i=1}^m (y^i + K)$ 是 K -有界和 K -闭的 (见文献[27]的定义 3.3.2). 据此, 由文献[27]的定理 3.3.6 可知 $\{y^1, \dots, y^m\}$ 的 K -界集是存在的.

定义 2.4.4 设 $S \subset \mathscr{X}$ 是非空凸集, $\varphi: S \rightarrow \mathscr{Y}$ 是映射.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$K\text{-bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K \\ \forall x^1, x^2 \in S, \quad (2.4.14)$$

则称 φ 是集合 S 上的 K -拟凸映射, 或映射 φ 在 S 上是 K -拟凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$K\text{-bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset \text{int}K(K \setminus \{\theta\}) \\ \forall x^1, x^2 \in S, \varphi(x^1) \neq \varphi(x^2), \quad (2.4.15)$$

则称 φ 是集合 S 上的 K -严格拟凸映射 (K -次严格拟凸映射), 或映射 φ 在 S 上是 K -严格拟凸的 (K -次严格拟凸的).

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$K\text{-bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset \text{int}K \\ \forall x^1, x^2 \in S, x^1 \neq x^2, \quad (2.4.16)$$

则称 φ 是集合 S 上的 K -强拟凸映射, 或映射 φ 在 S 上是 K -强拟凸的.

注 2.4.7 从定义 2.4.4 直接可以得知, 凸集上的 K -强拟凸映射也是 K -拟凸的、 K -严格拟凸的和 K -次严格拟凸的, 又 K -严格拟凸映射也是 K -次严格拟凸的, 反之不然. 然而, K -严格拟凸映射却不一定是 K -拟凸映射.

注 2.4.8 在定义 2.4.4 中当 $\mathscr{Y} = R$ 和 $K = R_+$ 时, $\varphi = f: S \rightarrow \mathscr{Y}$ 是实值函数. 这时, 因为

$$\bigcap_{i=1}^2 (f(x^i) + R_+) = [\max\{f(x^1), f(x^2)\}, +\infty),$$

由定义 2.4.3 可知

$$R_+ - \text{bou}\{f(x^1), f(x^2)\} = \max\{f(x^1), f(x^2)\}.$$

于是, 定义 2.4.4 中的 (2.4.14) 即

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \max\{f(x^1), f(x^2)\}$$

$$\forall x^1, x^2 \in S.$$

对照定义 2.3.1 中的 (2.3.1), 便知 R_+ -拟凸映射即拟凸函数. 同理, R_+ -严格拟凸映射和 R_+ -次严格拟凸映射 (这时二者等同) 即严格拟凸函数, R_+ -强拟凸映射即强拟凸函数.

注 2.4.9 在定义 2.4.4 中当 $\mathscr{Y} = R^m$ 和 $K = R_+^m$ ($m \geq 2$) 时, $\varphi = f: S \rightarrow R^m$ 是向量函数. 设 $f(x^i) = (f_1(x^i), \dots, f_m(x^i))^T$ ($i = 1, 2$), 则

$$\bigcap_{i=1}^2 (f(x^i) + R_+^m) = \begin{bmatrix} \max\{f_1(x^1), f_1(x^2)\} \\ \vdots \\ \max\{f_m(x^1), f_m(x^2)\} \end{bmatrix} + R_+^m,$$

于是由定义 2.4.3 可得知

$$R_+^m - \text{bou}\{f(x^1), f(x^2)\} = \begin{bmatrix} \max\{f_1(x^1), f_1(x^2)\} \\ \vdots \\ \max\{f_m(x^1), f_m(x^2)\} \end{bmatrix}.$$

据此,由定义 2.4.4 的(1)可知, f 是 S 上的 R_+^m -拟凸映射即对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$f_i(\lambda x^1 + (1 - \lambda)x^2) \leq \max\{f_i(x^1), f_i(x^2)\}, i = 1, \dots, m.$$

对照(2.3.1)得知,这时 f 的每一分量函数均是 S 上的拟凸函数. 同理, f 是 R_+^m -严格拟凸映射和 R_+^m -强拟凸映射也即它们的分量函数分别都是严格拟凸函数和强拟凸函数,而 f 是 R_+^m -次严格拟凸映射即它的每一分量都是拟凸函数并且其中至少有一个是严格拟凸函数.

锥凸映射类和锥拟凸映射类有如下相应关系.

定理 2.4.5 设 $S \subset \mathcal{V}$ 是非空凸集, $\varphi: S \rightarrow \mathcal{U}$ 是映射.

(1) 若 φ 是 S 上的 K -凸映射,则 φ 是 S 上的 K -拟凸映射.

(2) 若 φ 是 S 上的 K -严格凸映射(K -次严格凸映射),则 φ 是 S 上的 K -严格拟凸映射(K -次严格拟凸映射).

(3) 若 φ 是 S 上的 K -强凸映射,则 φ 是 S 上的 K -强拟凸映射.

证明 (1) 由于 φ 在 S 上是 K -凸的,按定义 2.4.1 的(1)可知,对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K,$$

或即

$$\varphi(\lambda x^1 + (1 - \lambda)x^2) \in \lambda\varphi(x^1) + (1 - \lambda)\varphi(x^2) - K. \quad (2.4.17)$$

依据定义 2.4.3,对任何 $\bar{y} \in K - \text{bou}\{\varphi(x^1), \varphi(x^2)\}$ 有 $\bar{y} \in \bigcap_{i=1}^2 (\varphi(x^i) + K)$, 因而

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} \subset \varphi(x^i) + K, i = 1, 2,$$

故

$$\varphi(x^i) \in K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - K, i = 1, 2.$$

据此,由(2.4.17)得

$$\begin{aligned}
\varphi(\lambda x^1 + (1 - \lambda)x^2) &\in \lambda(K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - K) \\
&\quad + (1 - \lambda)(K - \text{bou}\{\varphi(x^1), \\
&\quad \varphi(x^2)\} - K) - K \\
&= K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - K - K \\
&\subset K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - K,
\end{aligned}$$

从而

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K.$$

因此,由定义 2.4.4 的(1),即知 φ 在 S 上是 K -拟凸映射.

(2)和(3)与(1)的证明类似. \square

以下给出有关锥拟凸映射的两个基本结果.

定理 2.4.6 设 $S \subset \mathscr{V}$ 是非空凸集,则 $\varphi: S \rightarrow \mathscr{W}$ 是 K -拟凸映射当且仅当对任何正整数 $m \geq 2$ 和任意的 $\lambda_i \geq 0$ ($i = 1, \dots, m$), $\sum_{i=1}^m \lambda_i = 1$, 有

$$\begin{aligned}
K - \text{bou}\{\varphi(x^1), \dots, \varphi(x^m)\} - \varphi\left(\sum_{i=1}^m \lambda_i x^i\right) &\subset K \\
\forall x^1, \dots, x^m &\in S.
\end{aligned} \tag{2.4.18}$$

证明 充分性. 在(2.4.18)中取 $m = 2$, 由定义 2.4.4 中的(1)可知 φ 是 S 上的 K -拟凸映射.

必要性. 用数学归纳法证明, 当 $m = 2$ 时, 从(2.4.14)知(2.4.18)成立. 现设 $m \leq k$ ($k \geq 2$ 是正整数)时(2.4.18)成立, 以下证明(2.4.18)对于 $m = k + 1$ 成立. 不失一般性, 设 $\lambda_i > 0$ ($i = 1, \dots, k + 1$), $\sum_{i=1}^{k+1} \lambda_i = 1$ (若有某个 $\lambda_i = 0$, 则 $m \leq k$, 由假设(2.4.18)成立), 则有 $1 - \lambda_{k+1} = \sum_{i=1}^k \lambda_i > 0$. 记 $\sum_{i=1}^k \lambda_i x^i = (1 -$

$\lambda_{k+1})x$, 由 $x^i \in S$ 和 S 是凸集, 有 $x = \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \in S$. 由于已知 φ 在 S 上是 K -拟凸的, 按定义 2.4.4 的(1)得知

$$K + \text{bou}\{\varphi(x), \varphi(x^{k+1})\} = \varphi(\lambda_{k+1}x^{k+1} + (1 - \lambda_{k+1})x) \subset K,$$

或即

$$K + \text{bou}\left\{\varphi\left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i\right), \varphi(x^{k+1})\right\} \subset \varphi\left(\sum_{i=1}^{k+1} \lambda_i x^i\right) + K. \quad (2.4.19)$$

另外, 由 $\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} = 1$ 和归纳法假设(2.4.18)对于 $m = k$ 成立, 因而

$$K + \text{bou}\{\varphi(x^1), \dots, \varphi(x^k)\} = \varphi\left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i\right) \subset K. \quad (2.4.20)$$

因为显然有 $\bigcap_{i=1}^{k+1} (\varphi(x^i) + K) \subset \bigcap_{i=1}^k (\varphi(x^i) + K)$, 所以由定义 2.4.3 可以推知

$$\begin{aligned} & K + \text{bou}\{\varphi(x^1), \dots, \varphi(x^k), \varphi(x^{k+1})\} \\ & \subset K + \text{bou}\{\varphi(x^1), \dots, \varphi(x^k)\} + K. \end{aligned}$$

利用(2.4.20)并注意 K 是凸锥, 得到

$$K + \text{bou}\{\varphi(x^1), \dots, \varphi(x^{k+1})\} \subset \varphi\left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i\right) + K.$$

由定义 2.4.3 知

$$\begin{aligned} K + \text{bou}\{\varphi(x^1), \dots, \varphi(x^{k+1})\} & \subset \bigcap_{i=1}^{k+1} (\varphi(x^i) + K) \\ & \subset \varphi(x^{k+1}) + K, \end{aligned}$$

从而

$$\begin{aligned} K - \text{bou} \{ \varphi(x^1), \dots, \varphi(x^{k+1}) \} \\ \subset \left(\varphi \left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \right) + K \right) \cap (\varphi(x^{k+1}) + K). \end{aligned}$$

又从定义 2.4.3 易知

$$\begin{aligned} & \left(\varphi \left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \right) + K \right) \cap (\varphi(x^{k+1}) + K) \\ & \subset K - \text{bou} \left\{ \varphi \left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \right), \varphi(x^{k+1}) \right\} + K, \end{aligned}$$

因此有

$$\begin{aligned} K - \text{bou} \{ \varphi(x^1), \dots, \varphi(x^{k+1}) \} \\ \subset K - \text{bou} \left\{ \varphi \left(\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \right), \varphi(x^{k+1}) \right\} + K. \end{aligned}$$

最后,由(2.4.19)并注意到 K 是凸锥,即得

$$K - \text{bou} \{ \varphi(x^1), \dots, \varphi(x^{k+1}) \} \subset \varphi \left(\sum_{i=1}^{k+1} \lambda_i x^i \right) + K,$$

从而(2.4.18)对于 $m = k + 1$ 成立. \square

定理 2.4.7 设 $S \subset \mathcal{V}$ 是非空凸集,则 $\varphi: S \rightarrow \mathcal{W}$ 是 K -拟凸映射当且仅当对任意的 $c \in \mathcal{W}$, K -水平集 $H_S(\varphi, c)_K = \{x \in S | c - \varphi(x) \in K\}$ 是凸集.

证明 必要性. 对任意的 $c \in \mathcal{W}$, 任取 $x^1, x^2 \in H_S(\varphi, c)_K \subset S$, 则有 $c - \varphi(x^1) \in K$ 和 $c - \varphi(x^2) \in K$, 从而

$$c \in (\varphi(x^1) + K) \cap (\varphi(x^2) + K).$$

据此,由定义 2.4.3 得知,存在 $\bar{y} \in K - \text{bou} \{ \varphi(x^1), \varphi(x^2) \}$ 使 $c \in \bar{y} + K$, 或

$$c - \bar{y} \in K. \quad (2.4.21)$$

由于 φ 在 S 上是 K -拟凸的, 由定义 2.4.4 的(1)知, 对任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K,$$

从而有

$$\bar{y} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K.$$

于是, 注意到(2.4.21)和 K 是凸锥而得到

$$\begin{aligned} c - \varphi(\lambda x^1 + (1 - \lambda)x^2) \\ = (c - \bar{y}) + (\bar{y} - \varphi(\lambda x^1 + (1 - \lambda)x^2)) \in K + K \subset K, \end{aligned}$$

因而 $\lambda x^1 + (1 - \lambda)x^2 \in H_S(\varphi, c)_K$. 由此, 按定义 1.1.1 知 $H_S(\varphi, c)_K$ 是凸集.

充分性. 对任意的 $x^1, x^2 \in S$, 任取 $c \in K - \text{bou}\{\varphi(x^1), \varphi(x^2)\}$. 由于显然有 $K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} \subset (\varphi(x^1) + K) \cap (\varphi(x^2) + K)$, 故有

$$c - \varphi(x^1) \in K, c - \varphi(x^2) \in K,$$

于是 $x^1, x^2 \in H_S(\varphi, c)_K$. 因为 $H_S(\varphi, c)_K$ 是凸集, 故对任意的 $\lambda \in (0, 1)$ 有 $\lambda x^1 + (1 - \lambda)x^2 \in H_S(\varphi, c)_K$, 从而

$$c - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K.$$

由 $c \in K - \text{bou}\{\varphi(x^1), \varphi(x^2)\}$ 的任意性, 得到

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K,$$

因此由定义 2.4.4 的(1)知 φ 在 S 上是 K -拟凸的. \square

定理 2.1.3 和定理 2.3.3 曾经分别给出了在下半连续的条件
下, 严格凸函数是凸函数和严格拟凸函数是拟凸函数的结果. 以下
讨论 K -严格拟凸映射和 K -拟凸映射的类似关系.

设 $a \in \mathcal{S}$, $\omega \in \text{int}K$, 考虑 Minkowski 泛函

$$\mu_a: \mathscr{Y} \times \text{int}K \rightarrow R, (y, \omega) \mapsto \mu_a(y, \omega),$$

其中

$$\begin{aligned} \mu_a(y, \omega) &= \inf \{ \alpha \in R \mid y \in a + \alpha\omega - K \}, \\ y &\in \mathscr{Y}, \omega \in \text{int}K. \end{aligned} \quad (2.4.22)$$

显然, 对于任一 $y \in \mathscr{Y}$, $\mu_a(y, \omega)$ 是一确定的值, 并且有

$$y \in a + \mu_a(y, \omega)\omega - K. \quad (2.4.23)$$

引理 2.4.8 设 $y^1, y^2 \in \mathscr{Y}$.

(1) 若 $y^1 - y^2 \in K$, 则 $\mu_a(y^1, \omega) \geq \mu_a(y^2, \omega)$.

(2) 若 $y^1 - y^2 \in \text{int}K$, 则 $\mu_a(y^1, \omega) > \mu_a(y^2, \omega)$.

证明 (1) 由 $y^1 - y^2 \in K$, 则 $y^2 \in y^1 - K$. 对 y^1 利用 (2.4.23) 有

$$\begin{aligned} y^2 &\in (a + \mu_a(y^1, \omega)\omega - K) - K \\ &\subset a + \mu_a(y^1, \omega)\omega - K. \end{aligned}$$

据此, 由 (2.4.22) 得到 $\mu_a(y^2, \omega) \leq \mu_a(y^1, \omega)$.

(2) 由 $y^1 - y^2 \in \text{int}K$, 对 $\omega \in \text{int}K$ 存在 $\varepsilon > 0$, 使 $y^1 - y^2 - \varepsilon\omega \in K$, 即 $y^2 \in y^1 - \varepsilon\omega - K$. 利用 (2.4.23), 有

$$\begin{aligned} y^2 &\in (a + \mu_a(y^1, \omega)\omega - K) - \varepsilon\omega - K \\ &\subset a + (\mu_a(y^1, \omega) - \varepsilon)\omega - K. \end{aligned}$$

由此, 根据 (2.4.22) 得到 $\mu_a(y^2, \omega) \leq \mu_a(y^1, \omega) - \varepsilon < \mu_a(y^1, \omega)$. \square

引理 2.4.9 设 $a \in \mathscr{Y}$, 则对任意的 $y^1, y^2 \in \mathscr{Y}$ 有

$$\begin{aligned} &\inf \{ \mu_a(y, \omega) \mid y \in K - \text{bou}\{y^1, y^2\} \} \\ &= \max \{ \mu_a(y^1, \omega), \mu_a(y^2, \omega) \}, \end{aligned}$$

并且下确界可达.

证明 任取 $y \in K - \text{bou}\{y^1, y^2\}$, 由定义 2.4.3 可知 $y \in$

$(y^1 + K) \cap (y^2 + K)$, 从而

$$y - y^1 \in K, \quad y - y^2 \in K.$$

据此, 由引理 2.4.8 的(1), 有

$$\mu_a(y, \omega) \geq \mu_a(y^1, \omega), \quad \mu_a(y, \omega) \geq \mu_a(y^2, \omega),$$

于是

$$\mu_a(y, \omega) \geq \max \{ \mu_a(y^1, \omega), \mu_a(y^2, \omega) \},$$

从而得

$$\begin{aligned} \inf \{ \mu_a(y, \omega) \mid y \in K - \text{bou} \{ y^1, y^2 \} \} \\ \geq \max \{ \mu_a(y^1, \omega), \mu_a(y^2, \omega) \}. \end{aligned} \quad (2.4.24)$$

另一方面, 不妨设 $\mu_a(y^1, \omega) \geq \mu_a(y^2, \omega)$, 则有

$$y^1 \in a + \mu_a(y^1, \omega)\omega - K$$

和

$$y^2 \in a + \mu_a(y^2, \omega)\omega - K \subset a + \mu_a(y^1, \omega)\omega - K.$$

从而

$$a + \mu_a(y^1, \omega)\omega \in (y^1 + K) \cap (y^2 + K).$$

据此, 由定义 2.4.3 可知存在 $\bar{y} \in K - \text{bou} \{ y^1, y^2 \}$, 使 $a + \mu_a(y^1, \omega)\omega \in \bar{y} + K$, 也即

$$\bar{y} \in a + \mu_a(y^1, \omega)\omega - K.$$

根据(2.4.22)知

$$\mu_a(\bar{y}, \omega) \leq \mu_a(y^1, \omega) = \max \{ \mu_a(y^1, \omega), \mu_a(y^2, \omega) \},$$

再由(2.4.24)有

$$\begin{aligned} \inf \{ \mu_a(y, \omega) \mid y \in K - \text{bou} \{ y^1, y^2 \} \} &\leq \mu_a(\bar{y}, \omega) \\ &\leq \max \{ \mu_a(y^1, \omega), \mu_a(y^2, \omega) \} \\ &\leq \inf \{ \mu_a(y, \omega) \mid y \in K - \text{bou} \{ y^1, y^2 \} \}, \end{aligned}$$

因而得到

$$\begin{aligned} & \inf \{ \mu_a(y, \omega) \mid y \in K - \text{bou}\{y^1, y^2\} \} \\ &= \max \{ \mu_a(y^1, \omega), \mu_a(y^2, \omega) \}. \end{aligned}$$

显然, 其下确界在 $\bar{y} \in K - \text{bou}\{y^1, y^2\}$ 处达到. \square

设集合 $S \subset \mathscr{X}$, 考虑由映射 $\varphi: S \rightarrow \mathscr{Y}$ 确定的实值函数(泛函) $f_a: S \rightarrow R$:

$$f_a(x) = \mu_a(\varphi(x), \omega), \quad x \in S. \quad (2.4.25)$$

引理 2.4.10 设集合 $S \subset \mathscr{X}$ 非空, 则 $\varphi: S \rightarrow \mathscr{Y}$ 是 K -半连续的当且仅当对任意的 $a \in \mathscr{Y}$, $f_a: S \rightarrow R$ 是下半连续的.

证明 任取 $c \in R$, 由(2.4.22)可知对任意的 $a \in \mathscr{Y}$, $\mu_a(y, \omega) \leq c$ 意味着 $y \in a + c\omega - K$. 因此, 由(2.4.25)有

$$\{x \in S \mid f_a(x) \leq c\} = \{x \in S \mid \varphi(x) \in a + c\omega - K\}. \quad (2.4.26)$$

设 φ 在 S 上是 K -半连续的, 由定义 2.4.2 知(2.4.26)右端的集合是闭集. 据此, (2.4.26)左端的集合是闭集, 由定理 2.2.5 的(1)和(3)即知 f_a 是下半连续的. 反之, 若对任意的 $a \in \mathscr{Y}$, f_a 是下半连续的, 由定理 2.2.5 的(1)和(3)知 $\{x \in S \mid f_a(x) \leq 0\}$ 是闭集. 再根据(2.4.26), $\{x \in S \mid a - \varphi(x) \in K\}$ 是闭集, 于是由定义 2.4.2 得知 φ 是 K -半连续的. \square

引理 2.4.11 设集合 $S \subset \mathscr{X}$ 非空, $\varphi: S \rightarrow \mathscr{Y}$ 是映射.

(1) φ 在 S 上是 K -拟凸的当且仅当对任意的 $a \in \mathscr{Y}$, f_a 在 S 上是拟凸的.

(2) 若 φ 在 S 上是 K -严格拟凸的, 则对任意的 $a \in \mathscr{Y}$, f_a 在 S 上是严格拟凸的.

证明 (1) 必要性. 从 φ 在 S 上是 K -拟凸的, 由定义 2.4.4 的(1)可知, 对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou} \{ \varphi(x^1), \varphi(x^2) \} = \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K, \quad (2.4.27)$$

任取 $a \in \mathscr{A}$, 由引理 2.4.9 和 (2.4.25) 知存在 $\bar{y} \in K - \text{bou} \{ \varphi(x^1), \varphi(x^2) \}$, 使

$$\mu_a(\bar{y}, \omega) = \max \{ f_a(x^1), f_a(x^2) \}. \quad (2.4.28)$$

根据 $\bar{y} \in K - \text{bou} \{ \varphi(x^1), \varphi(x^2) \}$ 和 (2.4.27), 则 $\bar{y} = \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K$, 由引理 2.4.8 的(1)知

$$\begin{aligned} \mu_a(\bar{y}, \omega) &\geq \mu_a(\varphi(\lambda x^1 + (1 - \lambda)x^2), \omega) \\ &= f_a(\lambda x^1 + (1 - \lambda)x^2). \end{aligned}$$

于是, 由 (2.4.28) 得到

$$f_a(\lambda x^1 + (1 - \lambda)x^2) \leq \max \{ f_a(x^1), f_a(x^2) \},$$

从而按定义 2.3.1 的(1)即知 f_a 在 S 上是拟凸的.

充分性. 反之, 假设 φ 在 S 上不是 K -拟凸的, 根据定义 2.4.4 的(1), 则存在 $\bar{x}^1, \bar{x}^2 \in S$ 和 $\bar{\lambda} \in (0, 1)$ 使

$$K - \text{bou} \{ \varphi(\bar{x}^1), \varphi(\bar{x}^2) \} = \varphi(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2) \not\subset K.$$

于是, 有 $a^0 \in K - \text{bou} \{ \varphi(\bar{x}^1), \varphi(\bar{x}^2) \}$ 使 $a^0 = \varphi(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2) \notin K$, 即 $\varphi(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2) \notin a^0 - K$. 据此, 由 (2.4.22) 推知 $\mu_{a^0}(\varphi(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2), \omega) > 0$, 再按 (2.4.25) 得

$$f_{a^0}(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2) > 0. \quad (2.4.29)$$

另外, 由定义 2.4.3 知 $a^0 \in K - \text{bou} \{ \varphi(\bar{x}^1), \varphi(\bar{x}^2) \} \subset (\varphi(\bar{x}^1) + K) \cap (\varphi(\bar{x}^2) + K)$, 故有

$$\varphi(\bar{x}^1) \in a^0 - K, \quad \varphi(\bar{x}^2) \in a^0 - K.$$

根据 (2.4.22) 和 (2.4.25) 可知

$$f_{\bullet}^{\circ}(\bar{x}^1) = \mu_{\bullet}^{\circ}(\varphi(\bar{x}^1), \omega) \leq 0,$$

$$f_{\bullet}^{\circ}(\bar{x}^2) = \mu_{\bullet}^{\circ}(\varphi(\bar{x}^2), \omega) \leq 0.$$

因此,由(2.4.29)得到

$$f_{\bullet}^{\circ}(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2) > 0 \geq \max\{f_{\bullet}^{\circ}(\bar{x}^1), f_{\bullet}^{\circ}(\bar{x}^2)\},$$

根据定义 2.3.1 的(1),导致与对任意的 $a \in \mathscr{A}$, f_{\bullet} 在 S 上是拟凸的相矛盾.

(2)的证明与(1)的必要性证明部分类似,由定义 2.4.4 的(2)和定义 2.3.1 的(2),利用引理 2.4.9 和引理 2.4.8 的(2)可以推证. \square

定理 2.4.12 设 $S \subset \mathscr{X}$ 是非空凸集.若 $\varphi: S \rightarrow \mathscr{A}$ 是 K -半连续的 K -严格拟凸映射,则 φ 是 S 上的 K -拟凸映射.

证明 对于任意的 $a \in \mathscr{A}$, 因为 φ 在 S 上是 K -半连续的,由引理 2.4.10 知对应的 f_{\bullet} 是下半连续的.又根据 φ 在 S 上是 K -严格拟凸的,由引理 2.4.11 的(2), f_{\bullet} 在 S 上是严格拟凸的.现由 f_{\bullet} 是下半连续的严格拟凸函数,根据定理 2.3.3 得知 f_{\bullet} 是拟凸函数,因此由引理 2.4.11 的(1)即知 φ 在 S 上是 K -拟凸的. \square

以下讨论锥拟凸映射的几种条件.它们都是拟凸函数条件的推广,因而有的文献中也将它们作为锥拟凸映射的定义.

定理 2.4.13 设 $S \subset \mathscr{X}$ 是非空凸集, $\varphi: S \rightarrow \mathscr{A}$ 是映射.若对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 有

$$\varphi(x^1) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K$$

或

$$\varphi(x^2) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K \quad (2.4.30)$$

(Ferro 的恰当 K -拟凸映射,见文献[22]),则 φ 在 S 上是 K -拟凸的.

证明 不妨设有 $\varphi(x^1) - \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K$, 则

$$\varphi(x^1) \in \varphi(\lambda x^1 + (1 - \lambda)x^2) + K.$$

据定义 2.4.3 知 $K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} \subset \varphi(x^1) + K$, 故由上式并注意到 K 是凸锥而得到

$$\begin{aligned} K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} &\subset (\varphi(\lambda x^1 + (1 - \lambda)x^2) + K) + K \\ &\subset \varphi(\lambda x^1 + (1 - \lambda)x^2) + K, \end{aligned}$$

即

$$K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} = \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K.$$

因此, 由定义 2.4.4 的(1)即知 φ 在 S 上是 K -拟凸的. \square

注 2.4.10 在定理 2.4.13 中的 (2.4.30) 式中将 K 换作 $\text{int}K$ 或 $K \setminus \{0\}$, 并且补加条件 $\varphi(x^1) \neq \varphi(x^2)$, 则相应地可以得到 φ 在 S 上是 K -严格拟凸的或 K -次严格拟凸的. 若将 (2.4.30) 中的 K 换作 $\text{int}K$, 并加上条件 $x^1 \neq x^2$, 则相应地可得到 φ 在 S 上是 K -强拟凸的. 它们的证明与定理 2.4.13 的证明类似.

定理 2.4.14 设 $S \subset \mathcal{X}$ 是非空凸集, $\varphi: S \rightarrow \mathcal{Y}$ 是 K -拟凸映射当且仅当对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 有

$$\begin{aligned} y - \varphi(\lambda x^1 + (1 - \lambda)x^2) &\in K \\ \forall y \in \{y \in \mathcal{Y} \mid y - \varphi(x^1) &\in K, y - \varphi(x^2) \in K\} \end{aligned} \quad (2.4.31)$$

(Luc 的 K -拟凸映射, 见文献[23]).

证明 充分性. 条件 (2.4.31) 即

$$\begin{aligned} y &\in \varphi(\lambda x^1 + (1 - \lambda)x^2) + K \\ \forall y &\in (\varphi(x^1) + K) \cap (\varphi(x^2) + K). \end{aligned}$$

由定义 2.4.3 知

$$K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} \subset (\varphi(x^1) + K) \cap (\varphi(x^2) + K),$$

故得 $K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} \subset \varphi(\lambda x^1 + (1 - \lambda)x^2) + K$, 或即

$$K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} = \varphi(\lambda x^1 + (1 - \lambda)x^2) \subset K.$$

因此, 由定义 2.4.4 中的(1)得知 φ 在 S 上是 K -拟凸的.

必要性. 设 φ 在 S 上是 K -拟凸的, 则对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有(2.4.14), 也即

$$K = \text{bou}\{\varphi(x^1), \varphi(x^2)\} \subset \varphi(\lambda x^1 + (1 - \lambda)x^2) + K. \quad (2.4.32)$$

任取 $\bar{y} \in \{y \in \mathscr{U} \mid y - \varphi(x^1) \in K, y - \varphi(x^2) \in K\}$, 即

$$\bar{y} \in (\varphi(x^1) + K) \cap (\varphi(x^2) + K),$$

由定义 2.4.3 可知存在 $\bar{y} \in K = \text{bou}\{\varphi(x^1), \varphi(x^2)\}$ 使 $\bar{y} \in \bar{y} + K$. 据此, 由于(2.4.32)和 K 是凸锥而有

$$\bar{y} \in (\varphi(\lambda x^1 + (1 - \lambda)x^2) + K) + K \subset \varphi(\lambda x^1 + (1 - \lambda)x^2) + K.$$

从而得到 $\bar{y} = \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K$. \square

注 2.4.11 将定理 2.4.14 的 $\bar{y} = \varphi(\lambda x^1 + (1 - \lambda)x^2) \in K$ 中的 K 分别换作 $\text{int}K$ (或 $K \setminus \{0\}$, 并加 $\varphi(x^1) \neq \varphi(x^2)$) 和 $\text{int}K$ (加 $x^1 \neq x^2$), 则相应地有 φ 在 S 上是 K -严格拟凸的 (或 K -次严格拟凸的) 和 K -强拟凸的. 其证明与定理 2.4.14 的证明类似.

设 \mathscr{U}^* 是 \mathscr{U} 的对偶空间, $K^* \subset \mathscr{U}^*$ 是 K 的对偶锥.

定理 2.4.15 设 $S \subset \mathscr{X}$ 是非空凸集, $\varphi: S \rightarrow \mathscr{U}$ 是映射. 若对任意的 $\lambda \in (0, 1)$ 和任意的 $\nu^* \in K^* \setminus \{0\}$ 有

$$\begin{aligned} & \langle \nu^*, \varphi(\lambda x^1 + (1 - \lambda)x^2) \rangle \\ & \leq \max \{ \langle \nu^*, \varphi(x^1) \rangle, \langle \nu^*, \varphi(x^2) \rangle \} \quad \forall x^1, x^2 \in S \end{aligned} \quad (2.4.33)$$

(胡毓达等的 K -拟凸映射, 见文献[24]), 则 φ 在 S 上是 K -拟

凸的.

证明 用反证法. 假设 φ 在 S 上不是 K -拟凸的, 则由定义 2.4.4 的(1)可知, 存在 $\bar{x}^1, \bar{x}^2 \in S$ 和 $\bar{\lambda} \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(\bar{x}^1), \varphi(\bar{x}^2)\} - \varphi(\bar{\lambda}\bar{x}^1 + (1 - \bar{\lambda})\bar{x}^2) \not\subset K.$$

于是, 可取 $y^0 \in K - \text{bou}\{\varphi(\bar{x}^1), \varphi(\bar{x}^2)\}$ 使

$$y^0 - \varphi(\bar{\lambda}\bar{x}^1 + (1 - \bar{\lambda})\bar{x}^2) \notin K.$$

由此, 单点集 $\{y^0 - \varphi(\bar{\lambda}\bar{x}^1 + (1 - \bar{\lambda})\bar{x}^2)\}$ 与闭凸集 K 是严格可分离的, 故由定理 1.3.7 的(2)得知存在 $\bar{\nu}^* \in \mathcal{U}^* \setminus \{\theta\}$ 使

$$\langle \bar{\nu}^*, y^0 - \varphi(\bar{\lambda}\bar{x}^1 + (1 - \bar{\lambda})\bar{x}^2) \rangle < \langle \bar{\nu}^*, y \rangle \quad \forall y \in K.$$

我们指出, 实际上 $\bar{\nu}^* \in K^* \setminus \{\theta\}$ (因为否则, 若 $\bar{\nu}^* \notin K^*$, 由定义 1.4.6 的(1)知, 对某一 $z \in K$ 有 $\langle \bar{\nu}^*, z \rangle < 0$, 从而对任意大的 $\lambda > 0$ 有 $\langle \bar{\nu}^*, \lambda z \rangle < 0$. 令 $y = \lambda z \in K$ 代入上式, 则得其左端是有限值, 而右端可任意小, 导致矛盾). 于是, 在上式中取 $y = \theta \in K$, 则得存在 $\bar{\nu}^* \in K^* \setminus \{\theta\}$ 使

$$\langle \bar{\nu}^*, y^0 \rangle < \langle \bar{\nu}^*, \varphi(\bar{\lambda}\bar{x}^1 + (1 - \bar{\lambda})\bar{x}^2) \rangle. \quad (2.4.34)$$

另外, 由定义 2.4.3, 有

$$y^0 \in K - \text{bou}\{\varphi(\bar{x}^1), \varphi(\bar{x}^2)\} \subset \{\varphi(\bar{x}^1) + K\} \cap \{\varphi(\bar{x}^2) + K\},$$

故知

$$y^0 - \varphi(\bar{x}^1) \in K, \quad y^0 - \varphi(\bar{x}^2) \in K.$$

因为 $\bar{\nu}^* \in K^* \setminus \{\theta\}$, 所以有

$$\langle \bar{\nu}^*, y^0 - \varphi(\bar{x}^1) \rangle \geq 0, \quad \langle \bar{\nu}^*, y^0 - \varphi(\bar{x}^2) \rangle \geq 0,$$

从而

$$\langle \bar{\nu}^*, y^0 \rangle \geq \max\{\langle \bar{\nu}^*, \varphi(\bar{x}^1) \rangle, \langle \bar{\nu}^*, \varphi(\bar{x}^2) \rangle\}.$$

利用(2.4.34), 得到

$$\begin{aligned} & \langle \bar{\nu}^*, \varphi(\bar{\lambda} \bar{x}^1 + (1 - \bar{\lambda}) \bar{x}^2) \rangle \\ & > \max \{ \langle \bar{\nu}^*, \varphi(\bar{x}^1) \rangle, \langle \bar{\nu}^*, \varphi(\bar{x}^2) \rangle \}, \end{aligned}$$

导致与已设条件(2.4.33)矛盾. \square

注 2.4.12 在定理 2.4.15 的(2.4.33)中将“ \leq ”换作“ $<$ ”, 并且补加条件 $\varphi(x^1) \neq \varphi(x^2)$ 或 $x^1 \neq x^2$, 则相应地可得 φ 在 S 上是 K -严格拟凸的或 K -强拟凸的, 证明与定理 2.4.15 的证明类似.

最后, 介绍锥伪凸映射类.

定义 2.4.5 设 $S \subset \mathcal{X}$ 是非空开凸集, 映射 $\varphi: S \rightarrow \mathcal{Y}$ 是 Gâteaux 可微的.

(1) 若有

$$\varphi(x^1) - \varphi(x^2) \in \text{int}K \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -\text{int}K \quad \forall x^1, x^2 \in S, \quad (2.4.35)$$

则称 φ 是集合 S 上的 K -伪凸映射, 或映射 φ 在 S 上是 K -伪凸的.

(2) 若有

$$\begin{aligned} \varphi(x^1) - \varphi(x^2) \in K & \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -\text{int}K \\ & \forall x^1, x^2 \in S, \varphi(x^1) \neq \varphi(x^2), \end{aligned} \quad (2.4.36)$$

则称 φ 是集合 S 上的 K -严格伪凸映射, 或映射 φ 在 S 上是 K -严格伪凸的.

(3) 若有

$$\begin{aligned} \varphi(x^1) - \varphi(x^2) \in K & \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -\text{int}K \\ & \forall x^1, x^2 \in S, x^1 \neq x^2, \end{aligned} \quad (2.4.37)$$

则称 φ 是集合 S 上的 K -强伪凸映射, 或映射 φ 在 S 上是 K -强伪凸的.

(4) 若有

$$\varphi(x^1) - \varphi(x^2) \in K \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -K \quad \forall x^1, x^2 \in S, \quad (2.4.38)$$

则称 φ 是集合 S 上的 K -次伪凸映射, 或映射 φ 在 S 上是 K -次伪凸的.

注 2.4.13 从定义 2.4.5 直接可知, 开凸集上的 K -强伪凸映射也是 K -伪凸的, K -严格伪凸的和 K -次伪凸的, 又 K -严格伪凸映射也是 K -伪凸的, 反之不然.

注 2.4.14 在定义 2.4.5 中, 若 $\mathscr{Y} = R$ 和 $K = R_+$, 则 $\varphi = f: S \rightarrow R$ 是实值函数, (2.4.35) 和 (2.4.36) 成为

$$f(x^1) > f(x^2) \Rightarrow f'_{x^1}(x^2 - x^1) < 0 \quad \forall x^1, x^2 \in S.$$

由此, 按定义 2.3.3 的 (1), f 在 S 上是 R_+ -伪凸的和 R_+ -严格伪凸的, 即 f 是 S 上的伪凸函数. 同理, 这时 (2.4.37) 成为 (2.3.12), (2.4.38) 成为 (2.3.13), 因此 f 在 S 上是 R_+ -强(次)伪凸的, 即 f 是 S 上的强(次)伪凸函数.

注 2.4.15 在定义 2.4.5 中当 $\mathscr{Y} = R^m$ 和 $K = R^m_+$ ($m \geq 2$) 时, $\varphi = f: S \rightarrow R^m$ 是向量函数. 这时, 不难得知: f 是 R^m_+ -伪凸映射即它的每一分量函数都是伪凸函数, f 是 R^m_+ -强伪凸映射即它的每一分量函数都是强伪凸函数, f 是 R^m_+ -次伪凸映射即它的每一分量函数都是次伪凸函数.

定理 2.4.16 设 $S \subset \mathscr{X}$ 是非空开凸集, 映射 $\varphi: S \rightarrow \mathscr{Y}$ 是 Gâteaux 可微的.

(1) 若 φ 是 S 上的 K -凸映射, 则 φ 是 S 上的 K -伪凸映射和 K -次伪凸映射.

(2) 若 φ 是 S 上的 K -严格凸映射, 则 φ 是 S 上的 K -严格伪凸映射.

(3) 若 φ 是 S 上的 K -强凸映射, 则 φ 是 S 上的 K -强伪凸映射.

证明 (1) 由 φ 在 S 上是 K -凸的, 根据定义 2.4.1 的 (1) 可知, 对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^2) + (1 - \lambda)\varphi(x^1) - \varphi(\lambda x^2 + (1 - \lambda)x^1) \in K,$$

也即

$$\varphi(x^1 + \lambda(x^2 - x^1)) - \varphi(x^1) \in -\lambda[\varphi(x^1) - \varphi(x^2)] - K.$$

现设 $\varphi(x^1) - \varphi(x^2) \in \text{int}K$, 因为 K 是凸锥, 故有

$$\begin{aligned} \frac{\varphi(x^1 + \lambda(x^2 - x^1)) - \varphi(x^1)}{\lambda} &\in -[\varphi(x^1) - \varphi(x^2)] - K \\ &\subset -\text{int}K - K = -\text{int}K. \end{aligned}$$

令 $\lambda \rightarrow 0$, 按定义 2.3.2, 得

$$\varphi'_{x^1}(x^2 - x^1) = \lim_{\lambda \rightarrow 0} \frac{\varphi(x^1 + \lambda(x^2 - x^1)) - \varphi(x^1)}{\lambda} \in -\text{int}K.$$

因此, 由定义 2.4.5 中的(1)即知 φ 在 S 上是 K -伪凸的. 若设 $\varphi(x^1) - \varphi(x^2) \in K$, 同理可推得 $\varphi'_{x^1}(x^2 - x^1) \in -K$, 故由定义 2.4.5 的(4)知 φ 在 S 上是 K -次伪凸的.

(2)和(3)与(1)的证明类似. \square

定理 2.4.17 设 $S \subset \mathcal{X}$ 是非空开凸集, 映射 $\varphi: S \rightarrow \mathcal{Y}$ 是 Gâteaux 可微的. 若 φ 是 S 上的 K -拟凸映射, 则 φ 是 S 上的 K -次伪凸映射.

证明 由 φ 在 S 上是 K -拟凸的, 按定义 2.4.4 中的(1)可知, 对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(x^2), \varphi(x^1)\} - \varphi(\lambda x^2 + (1 - \lambda)x^1) \subset K. \quad (2.4.39)$$

设 $\varphi(x^1) - \varphi(x^2) \in K$, 则 $\varphi(x^1) \in \varphi(x^2) + K$, 从而

$$(\varphi(x^1) + K) \cap (\varphi(x^2) + K) = \varphi(x^1) + K.$$

于是, 由定义 2.4.3 可得 $K - \text{bou}\{\varphi(x^2), \varphi(x^1)\} = \varphi(x^1)$. 据此, 由(2.4.39)知 $\varphi(x^1) - \varphi(\lambda x^2 + (1 - \lambda)x^1) \in K$, 从而有

$$\frac{\varphi(x^1 + \lambda(x^2 - x^1)) - \varphi(x^1)}{\lambda} \in -K.$$

令 $\lambda \rightarrow 0$, 由定义 2.3.2 得

$$\varphi'_{x^1}(x^2 - x^1) = \lim_{\lambda \rightarrow 0} \frac{\varphi(x^1 + \lambda(x^2 - x^1)) - \varphi(x^1)}{\lambda} \in -K.$$

因此, 由定义 2.4.5 的(4), 即知 φ 在 S 上是 K -次伪凸的. \square

定理 2.4.18 设 $S \subset \mathcal{X}$ 是非空开凸集, 映射 $\varphi: S \rightarrow \mathcal{Y}$ 是 Gâteaux 可微的. 若对任意的 $\lambda \in (0, 1)$ 和任意的 $\nu^* \in K^* \setminus \{\theta\}$, 有

$$\begin{aligned} \langle \nu^*, \varphi(x^1) \rangle &> \langle \nu^*, \varphi(x^2) \rangle \Rightarrow \\ \langle \nu^*, \varphi'_{x^1}(x^2 - x^1) \rangle &< 0 \quad \forall x^1, x^2 \in S, \end{aligned} \quad (2.4.40)$$

则 φ 在 S 上是 K -伪凸的.

证明 用反证法, 假设在已设条件下 φ 在 S 上不是 K -伪凸的, 则由定义 2.4.5 的(1)可知, 存在 $\bar{x}^1, \bar{x}^2 \in S$ 和 $\bar{\lambda} \in (0, 1)$, 有

$$\varphi(\bar{x}^1) - \varphi(\bar{x}^2) \in \text{int}K \Rightarrow \varphi'_{\bar{x}^1}(\bar{x}^2 - \bar{x}^1) \notin -\text{int}K. \quad (2.4.41)$$

上式右端表明单点集 $\{-\varphi'_{\bar{x}^1}(\bar{x}^2 - \bar{x}^1)\}$ 与开凸集 $\text{int}K$ 可以分离, 于是由定理 1.3.7 的(1)得知, 存在 $\nu^* \in \mathcal{Y}^* \setminus \{\theta\}$ 使

$$\langle \nu^*, -\varphi'_{\bar{x}^1}(\bar{x}^2 - \bar{x}^1) \rangle \leq \langle \nu^*, y \rangle \quad \forall y \in \text{int}K. \quad (2.4.42)$$

与定理 2.4.15 的证明中同理有 $\nu^* \in K^* \setminus \{\theta\}$, 将它与 (2.4.41) 的左端相乘, 由定理 1.4.11 的(1)知

$$\langle \nu^*, \varphi(\bar{x}^1) - \varphi(\bar{x}^2) \rangle > 0,$$

或

$$\langle \nu^*, \varphi(\bar{x}^1) \rangle > \langle \nu^*, \varphi(\bar{x}^2) \rangle. \quad (2.4.43)$$

另外, 对于 (2.4.42), 令 $y \rightarrow 0$, 得到

$$\langle \nu^*, \varphi'_{\bar{x}^1}(\bar{x}^2 - \bar{x}^1) \rangle \geq 0. \quad (2.4.44)$$

(2.4.43)和(2.4.44)与(2.4.40)相矛盾. \square

注 2.4.16 在定理 2.4.18 的(2.4.40)中,将左端的“ $>$ ”换作“ \geq ”并补加条件 $\varphi(x^1) \neq \varphi(x^2)$ 或 $x^1 \neq x^2$,相应地可以得到 φ 在 S 上是 K -严格伪凸的或 K -强伪凸的.若将(2.4.40)中的“ $>$ ”和“ $<$ ”分别换作“ \geq ”和“ \leq ”,则可得到 φ 在 S 上是 K -次伪凸的.它们的证明与定理 2.4.18 的证明类似.

第 3 章 凸函数的次微分

凸函数的次微分是凸分析的核心内容,同时也是非光滑分析的重要组成部分.

这一章研究凸函数在一点处的变化率或微分问题. 可以看到,在通常的意义下,凸函数不一定是可微的. 为此,本章引进更能刻画凸函数在一点处变化特征的次梯度和次微分概念,并且讨论它们的有关性质. 最后,介绍次微分在凸规划研究中的应用.

§ 3.1 凸函数的方向导数

本节先考察凸函数沿给定方向的变化率也即方向导数问题. 可以看到,凸函数变化的一个直观特征是:它在其有效域中的一点处沿任何单边方向的方向导数总是存在的,但一般地说,它在同一点处沿不同方向的方向导数却不一定是相同的.

设 \mathcal{X} 是线性拓扑空间.

定义 3.1.1 设 $f: \mathcal{X} \rightarrow \mathbb{R} \cup \{\pm \infty\}$ 是实值函数,点 $x^0 \in \text{dom} f$, $d \in \mathcal{X}$.

(1) 若极限

$$f'_+(x^0; d) = \lim_{t \rightarrow 0^+} \frac{f(x^0 + td) - f(x^0)}{t} \quad (3.1.1)$$

存在,则称它是函数 f 在点 x^0 处沿方向 d 的(单边)右方向导数. 若极限

$$f'_-(x^0; d) = \lim_{t \rightarrow 0^-} \frac{f(x^0 + td) - f(x^0)}{t} \quad (3.1.2)$$

存在, 则称它是函数 f 在点 x^0 处沿方向 d 的(单边)左方向导数.

(2) 若 $f'_+(x^0; d) = f'_-(x^0; d)$, 即极限

$$f'(x^0; d) = \lim_{t \rightarrow 0} \frac{f(x^0 + td) - f(x^0)}{t} \quad (3.1.3)$$

存在, 则称它是函数 f 在点 x^0 处沿方向 d 的方向导数.

注 3.1.1 由定义 3.1.1 的(3)和定义 2.3.2 可知, 若 f 在点 x^0 处沿任意方向 $d \in \mathcal{H}$ 的方向导数均存在并且相等, 则 f 在点 x^0 处是 Gâteaux 可微的, 映射 $f'_{x^0}: \mathcal{H} \rightarrow R, d \mapsto f'_{x^0}(d) = f'(x^0; d)$ 即是 f 在点 x^0 处的 Gâteaux 导数.

定理 3.1.1 设 $f: \mathcal{H} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $x \in \text{int}(\text{dom} f)$, 则对任意的 $d \in \mathcal{H}$, $f'_+(x; d)$ 和 $f'_-(x; d)$ 存在, 并且

$$f'(x; d) \leq f'_+(x; d).$$

证明 设 $t \in R$, 记 $g(t) = f(x + td)$. 因为 f 是正常凸的, 由定理 2.1.8 和定义 2.1.5 可知, $g(t)$ 是 R 上的正常凸函数, 并且 $g(0) = f(x)$ 为有限值. 据此, 可以推知 $\frac{g(t) - g(0)}{t}$ 是 $t > 0$ 的非减函数, 从而极限 $\lim_{t \rightarrow 0^+} \frac{g(t) - g(0)}{t}$ 存在. 由定义 3.1.1 的(1)即得

$$\begin{aligned} f'_+(x; d) &= \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} \\ &= \lim_{t \rightarrow 0^+} \frac{g(t) - g(0)}{t} \end{aligned}$$

存在. 同理, 可以推知 $f'_-(x; d)$ 存在, 并且易知有 $f'_-(x; d) \leq f'_+(x; d)$. \square

定理 3.1.1 指出正常凸函数在其有效域内的任一点处沿任何方向的单边方向导数总是存在的, 并且其左方向导数值不超过右方向导数值. 但是, 下面的例子说明, 凸函数在其有效域中的一点处其方向导数可能不存在, 因而由注 3.1.1 可知, 它在该点处不是

Gâteaux 可微的.

例 3.1.1 范数函数的方向导数. 考虑例 2.1.1 中的范数函数

$$f(x) = \|x\|, x \in R^n.$$

在例 2.1.1 中已指出它是 R^n 上的凸函数. 考察点 $\theta \in R^n$ 和任一单位向量 $d \in R^n$, $\|d\| = 1$. 由 (3.1.1) 和 (3.1.2) 可知, f 在点 θ 处沿 d 的(单边)右方向导数和左方向导数分别是

$$f'_+(\theta; d) = \lim_{t \rightarrow 0^+} \frac{\|\theta + td\| - \|\theta\|}{t} = 1$$

和

$$f'_-(\theta; d) = \lim_{t \rightarrow 0^-} \frac{\|\theta + td\| - \|\theta\|}{t} = -1.$$

由于 $f'_+(\theta; d) \neq f'_-(\theta; d)$, 故按定义 3.1.1 的(2), f 在点 θ 处沿任何方向 d 的方向导数均不存在.

凸函数的单边方向导数有以下性质.

定理 3.1.2 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $x \in \text{dom} f$, $d \in \mathcal{X}$, 则 $f'_+(x; d)$ 关于 d 是正常凸函数.

证明 对于任意的 $d^1, d^2 \in \mathcal{X}$ 和任意的 $\lambda \in (0, 1)$, 由 f 是凸函数, 我们有

$$\begin{aligned} & f(x + t(\lambda d^1 + (1 - \lambda)d^2)) - f(x) \\ &= f(\lambda(x + td^1) + (1 - \lambda)(x + td^2)) - f(x) \\ &\leq \lambda f(x + td^1) + (1 - \lambda)f(x + td^2) - f(x) \\ &= \lambda[f(x + td^1) - f(x)] + (1 - \lambda)[f(x + td^2) - f(x)]. \end{aligned}$$

两边除以 t , 并令 $t \rightarrow 0^+$, 得到

$$f'_+(x; \lambda d^1 + (1 - \lambda)d^2)$$

$$\begin{aligned}
&= \lim_{t \rightarrow 0^+} \frac{f(x + t(\lambda d^1 + (1 - \lambda)d^2)) - f(x)}{t} \\
&\leq \lambda \lim_{t \rightarrow 0^+} \frac{f(x + td^1) - f(x)}{t} \\
&\quad + (1 - \lambda) \lim_{t \rightarrow 0^+} \frac{f(x + td^2) - f(x)}{t} \\
&= \lambda f'_+(x; d^1) + (1 - \lambda)f'_+(x; d^2).
\end{aligned}$$

据此,由定义 2.1.1 的(1)即知 $f'_+(x; d)$ 关于 d 是凸函数. 又由 f 是正常凸函数,按定义 3.1.1 的(1),进而可知 $f'_+(x; d)$ 关于 d 是正常凸的. \square

定理 3.1.3 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $d \in \mathcal{X}$, f 在点 $x \in \text{dom} f$ 处是连续的.

(1) $f'_+(x; d)$ 关于 d 是连续的.

(2) $f'_+(x; d)$ 关于 d 是正齐次和次可加的(即次线性的).

证明 (1) 由定理 3.1.2 得知 $f'_+(x; d)$ 关于 d 是正常凸函数. 因此,为证明 $f'_+(x; d)$ 关于 d 是连续的,依据定理 2.2.1,只需证明它在某邻域内上有界. 事实上,从 f 在点 x 处是连续的可知,对于某零点邻域中的 d , $f(x + d)$ 上有界. 于是,由

$$f'_+(x; d) = \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} \leq f(x + d) - f(x)$$

即知 $f'_+(x; d)$ 在此零点邻域内也上有界.

(2) 对于 $\alpha \geq 0$, 由(3.1.1)有

$$\begin{aligned}
f'_+(x; \alpha d) &= \lim_{t \rightarrow 0^+} \frac{f(x + t\alpha d) - f(x)}{t} \\
&= \alpha \lim_{t \rightarrow 0^+} \frac{f(x + t\alpha d) - f(x)}{\alpha t} \\
&= \alpha f'_+(x; d),
\end{aligned}$$

故正齐次性成立. 又对任意的 $d^1, d^2 \in \mathcal{X}$, 因 f 是凸函数,故

$$\begin{aligned}
f'_+(x; d^1 + d^2) &= \lim_{t \rightarrow 0^+} \frac{f(x + t(d^1 + d^2)) - f(x)}{t} \\
&\leq \lim_{t \rightarrow 0^+} \frac{f(x + 2td^1) + f(x + 2td^2) - 2f(x)}{2t} \\
&= \lim_{t \rightarrow 0^+} \frac{f(x + 2td^1) - f(x)}{2t} + \lim_{t \rightarrow 0^+} \frac{f(x + 2td^2) - f(x)}{2t} \\
&= f'_+(x; d^1) + f'_+(x; d^2).
\end{aligned}$$

由此, $f'_+(x; d)$ 关于 d 具有次可加性. 综合以上结果, $f'_+(x; d)$ 关于 d 是次线性泛函. \square

对于 Gâteaux 可微的正常凸函数, 有如下结论.

设 \mathcal{X}^* 是 \mathcal{X} 的对偶空间.

定理 3.1.4 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $x \in \text{dom} f$. 若 f 在点 x 处是 Gâteaux 可微的, 则 $f'_x(d) = f'_+(x; d)$ 关于 d 是连续线性泛函, 即存在连续线性泛函 $x^* \in \mathcal{X}^*$, 使

$$f'_x(d) = \langle x^*, d \rangle \quad \forall d \in \mathcal{X}. \quad (3.1.4)$$

证明 因为 f 在点 x 处是 Gâteaux 可微的, 由定理 2.3.7, 则 f 在点 x 处是连续的. 因此, 根据定理 3.1.3 得知 $f'_+(x; d)$ 关于 d 是连续的次线性泛函.

又从 f 在点 x 处是 Gâteaux 可微的, 由定义 2.3.2 和定义 3.1.1 的(2)可知, 对任意的 $d \in \mathcal{X}$, 有 $f'_x(d) = f'_+(x; d) = f'_-(x; d) = -f'_-(x; -d)$. 于是, 由已知 $f'_x(d) = f'_+(x; d)$ 和 $-f'_x(d) = f'_-(x; -d)$ 关于 d 都是连续的次线性泛函, 即得 $f'_x(d)$ 关于 d 是连续线性泛函. \square

注 3.1.2 从定理 3.1.4 得知, 对于在某点 $x \in \mathcal{X}$ 处 Gâteaux 可微的正常凸函数, 它在该点处的 Gâteaux 导数 f'_x 与满足

$$f'_x(d) = \langle x^*, d \rangle \quad \forall d \in \mathcal{X}$$

的连续线性泛函 x^* 对应. 因而也可将此连续线性泛函定义为它在

该点处的 Gâteaux 导数.

注 3.1.3 在微积分中我们知道, 对于 Euclid 空间上的可微函数 $f: R^n \rightarrow R$, 设 $\nabla f(x)$ 是它在点 $x \in R^n$ 处的梯度, $f'(x; d)$ 是在点 x 处沿任一方向 d 的方向导数, 则有 $f'(x; d) = \nabla f(x)^T d$, 或即

$$f'_x(d) = \langle \nabla f(x), d \rangle \quad \forall d \in R^n.$$

将它与 (3.1.4) 当 $\mathcal{X} = R^n$ 时相比, 即知当 f 是正常凸函数时, 与它在点 x 处的 Gâteaux 导数对应的连续线性泛函即是梯度 $\nabla f(x)$. 可见, Gâteaux 导数是梯度的一种推广.

下面介绍另一可微的概念.

设 \mathcal{B} 是 Banach 空间, \mathcal{B}^* 是 \mathcal{B} 的对偶空间.

定义 3.1.2 设 $f: \mathcal{B} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 点 $x^0 \in \text{dom} f$, $y \in \mathcal{B}$. 若存在 $x' \in \mathcal{B}^*$ 使得

$$\lim_{y \rightarrow x^0} \frac{f(y) - f(x^0) - \langle x', y - x^0 \rangle}{\|y - x^0\|} = 0, \quad (3.1.5)$$

则称函数 f 在点 x^0 处是 Frechet 可微的. 称 x' 是 f 在点 x^0 处的 Frechet 导数, 记作 $f'(x^0) = x'$. 并且称 $\langle x', y - x^0 \rangle$ 是 f 在点 x^0 处的 Frechet 微分, 记作 $df(x^0) = \langle x', y - x^0 \rangle$.

注 3.1.4 显然, (3.1.5) 与

$$f(y) - f(x^0) = \langle x', y - x^0 \rangle + o(\|y - x^0\|), \quad y \in \mathcal{B} \quad (3.1.6)$$

等价, 其中 $\frac{o(\|y - x^0\|)}{\|y - x^0\|} \rightarrow 0 (y \rightarrow x^0)$. 对于 Euclid 空间上的可微函数 $f: R^n \rightarrow R$, 设 $df(x^0)$ 是它在点 $x^0 \in R^n$ 处的全微分, 则有

$$f(y) - f(x^0) = df(x^0) + o(\|y - x^0\|).$$

当 $\mathcal{B} = R^n$ 时将它与 (3.1.6) 相比, 即知这时 Frechet 微分即是全微分:

$$\langle x', y - x^0 \rangle = df(x^0).$$

可见, Frechet 微分是全微分的推广, 我们仍记它为 $df(x^0)$.

定理 3.1.5 设 $f: \mathcal{B} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, $x \in \text{dom} f$. 若 f 在点 x 处是 Frechet 可微的, 则 f 在点 x 处是 Gâteaux 可微的, 并且

$$f'_x(d) = \langle f'(x), d \rangle \quad \forall d \in \mathcal{B}.$$

证明 对任何 $d \in \mathcal{B}$, 令 $y = x + td$ (其中 $t \in R$). 由定义 3.1.1 的(2)和(3.1.5)有

$$\begin{aligned} f'(x; d) &= \lim_{t \rightarrow 0} \frac{f(x + td) - f(x)}{t} \\ &= \lim_{y \rightarrow x} \frac{f(y) - f(x) - \langle f'(x), y - x \rangle}{\|y - x\|} \|d\| \\ &\quad + \langle f'(x), d \rangle \\ &= \langle f'(x), d \rangle. \end{aligned}$$

因为 $d \in \mathcal{B}$ 是任意的, 所以由定义 2.3.2 和定义 3.1.1 的(2)即得 $f'_x(d) = \langle f'(x), d \rangle$. \square

推论 3.1.6 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $x \in \text{dom} f$. 若 f 在点 x 处是 Frechet 可微的, 则 $f'(x)$ 是 \mathcal{B} 上的连续线性泛函.

证明 由定理 3.1.5 和定理 3.1.4 可得. \square

§ 3.2 次梯度和次微分

从上一节我们知道, 凸函数在其有效域内虽然总存在单边方向导数, 但却不一定是 Gâteaux 可微的. 本节引入凸函数的次微分概念, 并且阐明在次微分的意义下, 凸函数在其连续点处总是次可微的.

设 \mathcal{X} 是线性拓扑空间, \mathcal{X}^* 是 \mathcal{X} 的对偶空间.

定义 3.2.1 设 $f: \mathcal{H} \rightarrow R \cup \{\pm \infty\}$ 是凸函数, 点 $x^0 \in \mathcal{H}$. 若 $x^* \in \mathcal{H}^*$, 并且

$$\langle x^*, y - x^0 \rangle \leq f(y) - f(x^0) \quad \forall y \in \mathcal{H}, \quad (3.2.1)$$

则称 x^* 是 f 在点 x^0 处的次梯度. f 在点 x^0 处的所有次梯度的集合称为 f 在点 x^0 处的次微分, 记作 $\partial f(x^0)$, 即

$$\begin{aligned} \partial f(x^0) = \{x^* \in \mathcal{H}^* \mid \langle x^*, y - x^0 \rangle \\ \leq f(y) - f(x^0) \quad \forall y \in \mathcal{H}\}. \end{aligned} \quad (3.2.2)$$

若 $\partial f(x^0) \neq \emptyset$, 则称 f 在点 x^0 处是次可微的. 若 f 在集合 $S \subset \mathcal{H}$ 的每一点处是次可微的, 则称 f 在 S 上是次可微的.

例 3.2.1 范数函数的次微分. 再考虑例 2.1.1 中的凸函数

$$f(x) = \|x\|, \quad x \in R^n.$$

由例 3.1.1 已知, 它在点 θ 处沿任何方向的方向导数均不存在. 但据定义 3.2.1, $x^* \in \mathcal{H}^* = R^n$ 是 f 在点 θ 处的次梯度, 意即对任意的 $y \in R^n$ 有 $\langle x^*, y \rangle \leq \|y\| - \|\theta\|$, 也即 $x^{*T} y \leq \|y\|$ 或 $\|x^*\| \leq 1$. 由此可知, R^n 中满足 $\|x^*\| \leq 1$ 的向量 x^* 都是 f 在点 θ 处的次梯度, 而 f 在点 θ 处的次微分即单位球

$$\partial f(\theta) = \{x^* \in R^n \mid \|x^*\| \leq 1\},$$

因而 f 在点 θ 处是次可微的. 特别地, 对于 $n = 1$ 时, 凸函数 $f(x) = |x|$ 在点 θ 处的次微分是闭区间 $\partial f(\theta) = [-1, 1]$, 其任一次梯度 $x^* \in [-1, 1]$ (图 3.2.1 中的 x^* 对应 $\zeta = |x|$ 在点 θ 处切线的斜率).

注 3.2.1 设 x^* 是 f 在点 $x^0 \in \mathcal{H}$ 处的次梯度. 令 $\beta = \langle x^*, x^0 \rangle - f(x^0)$, 考虑 \mathcal{H}

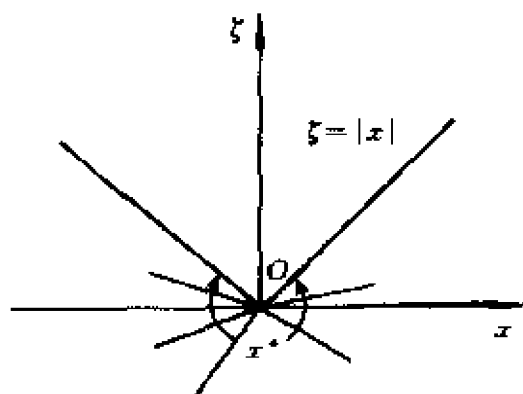


图 3.2.1

$\times R$ 中非垂直的闭超平面

$$\tilde{H} = \{(x, \zeta) \in \mathcal{X} \times R \mid \langle x^*, x \rangle - \zeta = \beta\}, \quad (3.2.3)$$

则有 $(x^0, f(x^0)) \in \tilde{H}$. 由于 f 是凸函数, 根据定理 2.1.5 知它的上图象 $\text{epi} f \subset \mathcal{X} \times R$ 是凸集, 并且显然有 $(x^0, f(x^0)) \in \text{epi} f$. 另外, 任取 $(y, \eta) \in \text{epi} f$, 由定义 2.1.6 的(1)和定义 3.2.1 得

$$\eta \geq f(y) \geq \langle x^*, y - x^0 \rangle + f(x^0),$$

即

$$\langle x^*, y \rangle - \eta \leq \langle x^*, x^0 \rangle - f(x^0),$$

于是

$$(y, \eta) \in \tilde{H}_- = \{(x, \zeta) \in \mathcal{X} \times R \mid \langle x^*, x \rangle - \zeta \leq \beta\}.$$

因为 $(y, \eta) \in \text{epi} f$ 是任取的, 故得 $\text{epi} f \subset \tilde{H}_-$. 根据以上的讨论, 由定义 1.3.3 得知 \tilde{H} 是 $\text{epi} f$ 在点 $(x^0, f(x^0))$ 处的非垂直的闭支撑超平面(图 3.2.2(a)). 若 f 在点 x^0 处的次梯度不唯一, 则 f 的上图象 $\text{epi} f$ 在点 $(x^0, f(x^0))$ 处的支撑超平面也不唯一(图 3.2.2(b)). 特别是, 由(3.2.3)当 $\mathcal{X} = R$ 时, 可知单变量凸函数 f 在点 x^0 处的次梯度 x^* 就是它在点 x^0 处支撑切线 \tilde{H} 的斜率.

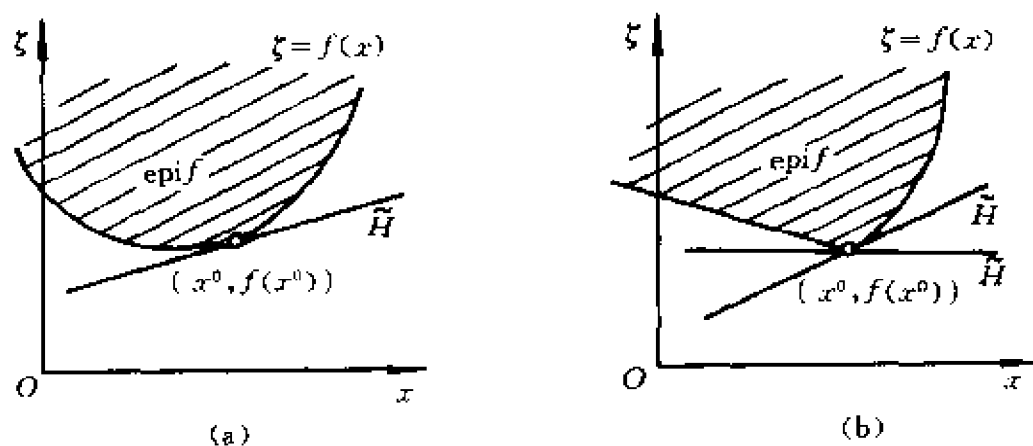


图 3.2.2

注 3.2.2 我们指出, 凸函数 f 在任意点 $x \in \mathcal{X}$ 处的次微分 $\partial f(x)$ 都是闭凸集. 事实上, 若 $\partial f(x)$ 是空集或单点集, 它显然是

闭凸的. 设 $\partial f(x)$ 至少包含两个点. 由 (3.2.2) 易知它是闭集. 此外, 任取 $x_1^*, x_2^* \in \partial f(x)$ 和任意的 $\lambda \in (0, 1)$, 由 (3.2.2) 有

$$\langle x_1^*, y - x \rangle \leq f(y) - f(x) \quad \forall y \in \mathcal{X}, \quad (3.2.4)$$

$$\langle x_2^*, y - x \rangle \leq f(y) - f(x) \quad \forall y \in \mathcal{X}. \quad (3.2.5)$$

作 $\lambda(3.2.4) + (1 - \lambda)(3.2.5)$, 则得

$$\begin{aligned} & \langle \lambda x_1^* + (1 - \lambda)x_2^*, y - x \rangle \\ & \leq [\lambda f(y) + (1 - \lambda)f(y)] - [\lambda f(x) + (1 - \lambda)f(x)] \\ & = f(y) - f(x). \end{aligned}$$

于是, 再由 (3.2.2) 即知 $\lambda x_1^* + (1 - \lambda)x_2^* \in \partial f(x)$, 从而 $\partial f(x)$ 是凸集. 此外, $\partial f(x)$ 还是弱*紧的 (见 [25] 定理 4.1.2 的推论 2).

下面是次微分的存在性定理.

定理 3.2.1 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数. 若 f 在某点 $x^0 \in \text{int}(\text{dom} f)$ 处是连续的, 则

$$\partial f(x) \neq \emptyset \quad \forall x \in \text{int}(\text{dom} f).$$

证明 首先, 因为 f 是凸函数, 故它的上图象 $\text{epi} f \subset \mathcal{X} \times R$ 是凸集. 其次, 由 f 在点 x^0 处是连续的, 可知存在 x^0 的邻域 $U(x^0)$ 和 $t > 0$ 使得

$$f(x) \leq t \quad \forall x \in U(x^0).$$

因此, 当 $x \in U(x^0)$ 和 $\eta \geq t$ 时, 有 $(x, \eta) \in \text{epi} f$, 从而 $\text{int}(\text{epi} f) \neq \emptyset$. 另外, 显然有 $(x^0, f(x^0)) \in \text{rb}(\text{epi} f)$, 于是根据推论 1.3.12 得知存在闭超平面 $H \subset \mathcal{X} \times R$, 它在点 $(x^0, f(x^0))$ 处正常支撑 $\text{epi} f$. 由于 $x^0 \in \text{int}(\text{dom} f)$, 则可推知 H 是非垂直的 (见 [15] 的引理 5.33). 现设

$$H = \{(x, \zeta) \in \mathcal{X} \times R \mid \langle x^*, x \rangle + \alpha \zeta = \beta\},$$

并且不失一般性让 $\text{epi} f \subset H_+$, 则对任意的 $(y, \eta) \in \text{epi} f$ 有

$$\langle x^*, y \rangle + \alpha \eta \geq \beta.$$

因为 η 可以充分大, 故 $\alpha > 0$. 由于特别是对 $y \in \text{dom} f$ 有 $(y, f(y)) \in \text{epi} f$, 从上式可知

$$\langle x^*, y \rangle + \alpha f(y) \geq \beta. \quad (3.2.6)$$

另外, 由 $(x^0, f(x^0))$ 是支撑点, 则 $(x^0, f(x^0)) \in H$, 因而有

$$\langle x^*, x^0 \rangle + \alpha f(x^0) = \beta.$$

将它代入 (3.2.6) 有 $\langle x^*, y \rangle + \alpha f(y) \geq \langle x^*, x^0 \rangle + \alpha f(x^0)$, 于是得到

$$\left\langle -\frac{1}{\alpha} x^*, y - x^0 \right\rangle \leq f(y) - f(x^0).$$

由于当 $y \notin \text{dom} f$ 时, 按定义 2.1.4 知上式显然也成立, 故由定义 3.2.1 得知 $-\frac{1}{\alpha} x^* \in \partial f(x^0)$, 从而 $\partial f(x^0) \neq \emptyset$.

由于 f 在点 $x^0 \in \text{int}(\text{dom} f)$ 处是连续的, 它在 $\text{int}(\text{dom} f)$ 内一点的某邻域内上有界, 故根据定理 2.2.1, f 在 $\text{int}(\text{dom} f)$ 内是连续的. 由此, 按照上面的证明得知, 对任意的 $x \in \text{int}(\text{dom} f)$ 有 $\partial f(x) \neq \emptyset$. \square

现在给出次梯度与方向导数之间的关系.

定理 3.2.2 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数. 若 f 在点 $x \in \text{dom} f$ 处是连续的, 则 $x^* \in \partial f(x)$ 当且仅当

$$\langle x^*, d \rangle \leq f_+(x; d) \quad \forall d \in \mathcal{X}. \quad (3.2.7)$$

证明 必要性. 设 $x^* \in \partial f(x)$, 由定义 3.2.1 可知, 对任意的 $d \in \mathcal{X}$ 和 $t > 0$ 有

$$\langle x^*, td \rangle \leq f(x + td) - f(x).$$

因此, 按定义 3.1.1 的(1)即得

$$\langle x^*, d \rangle \leq \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} = f'_+(x; d).$$

充分性. 对于任意的 $d \in \mathcal{H}$, 易知 $\frac{f(x + td) - f(x)}{t}$ 是 $t > 0$ 的非减函数. 据此, 由 (3.2.7) 得知有

$$\begin{aligned} \langle x^*, d \rangle &\leq f'_+(x; d) = \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} \\ &\leq \frac{f(x + td) - f(x)}{t}, \end{aligned}$$

从而 $\langle x^*, td \rangle \leq f(x + td) - f(x)$. 令 $y = x + td$, 则 $\langle x^*, y - x \rangle \leq f(y) - f(x)$, 于是由定义 3.2.1 即知 $x^* \in \partial f(x)$. \square

注 3.2.3 事实上, 在定理 3.2.2 的基础上, 还可以得到

$$f'_+(x; d) = \sup_{x^* \in \partial f(x)} \langle x^*, d \rangle \quad \forall d \in \mathcal{H}$$

(见 [25] 的定理 4.1.2).

以下定理表明, 在一定意义的可微条件下, 次微分即为该意义下的导数或梯度.

定理 3.2.3 设 $f: \mathcal{H} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, f 在点 $x \in \text{dom} f$ 处是连续的. 若 f 在点 x 处是 Gâteaux 可微的, 并且其 Gâteaux 导数是 $f'_x \in \mathcal{H}^*$, 则 $\partial f(x) = \{f'_x\}$.

证明 因为 f 在点 x 处是 Gâteaux 可微的, 故由定理 3.1.4 和注 3.1.2 有

$$f'_x(d) = \langle f'_x, d \rangle \quad \forall d \in \mathcal{H}. \quad (3.2.8)$$

由于按定义 2.3.2, 这时对任意的 $d \in \mathcal{H}$ 有 $f'_x(d) = f'(x; d)$, 因而也有

$$f'_x(d) = f'_+(x; d) \quad \forall d \in \mathcal{H}. \quad (3.2.9)$$

结合 (3.2.8), 知

$$\langle f'_x, d \rangle = f'_+(x; d) \quad \forall d \in \mathcal{H},$$

因此由定理 3.2.2 得到 $f'_x \in \partial f(x)$.

假设另有 $x^* \in \partial f(x)$, 则由定理 3.2.2 和 (3.2.9) 有

$$\langle x^*, d \rangle \leq f'_x(d) \quad \forall d \in \mathcal{X}.$$

据此, 注意 (3.2.8) 得

$$\langle f'_x - x^*, d \rangle \geq 0 \quad \forall d \in \mathcal{X}.$$

由 d 的任意性便导致 $x^* = f'_x$. \square

设 \mathcal{B} 是 Banach 空间.

定理 3.2.4 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, f 在点 $x \in \text{dom} f$ 处是连续的. 若 f 在点 x 处是 Frechet 可微的, 并且其 Frechet 导数是 $f'(x)$, 则 $\partial f(x) = \{f'(x)\}$.

证明 由 f 在点 x 处是 Frechet 可微的, 据定理 3.1.5, 它在该点处也是 Gâteaux 可微的, 并且

$$f'_x(d) = \langle f'(x), d \rangle \quad \forall d \in \mathcal{B}.$$

设 $f'_x \in \mathcal{X}^*$ 是 f 在点 x 处的 Gâteaux 导数, 由定理 3.1.4 和注 3.1.2 对比 (3.1.4) 和上式, 得知 $f'(x) = f'_x$. 由定理 3.2.3 即得 $\partial f(x) = \{f'(x)\}$. \square

推论 3.2.5 设 $f: R^n \rightarrow R$ 是正常凸函数. 若 f 在点 $x \in \text{dom} f$ 处是可微的, 则 $\partial f(x) = \{\nabla f(x)\}$.

证明 由注 3.1.3 和定理 3.2.3 可得证. \square

例 3.2.2 仿射函数的次微分. 设 $a \in \mathcal{X}$, $\beta \in R$, 从例 2.1.1 已知仿射函数

$$l(x) = \langle a, x \rangle + \beta, \quad x \in \mathcal{X}$$

是 \mathcal{X} 上的凸函数. 由定义 3.2.1, $x^* \in \mathcal{X}^*$ 是函数 l 在点 $x \in \mathcal{X}$ 处的次梯度当且仅当对任意的 $y \in \mathcal{X}$ 有

$$\langle x^*, y - x \rangle \leq \langle a, y \rangle + \beta - \langle a, x \rangle - \beta,$$

或即 $\langle a - x^*, y - x \rangle \geq 0$. 由 y 的任意性, 即知 $x^* = a$, 于是

$$\partial l(x) = \{a\}.$$

例 3.2.3 指示函数的次微分. 设 $S \subset \mathcal{X}$ 是非空凸集, 考虑例 2.1.1 中的指示函数

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathcal{X} \setminus S. \end{cases}$$

例 2.1.1 已指出它是 S 上的凸函数. 由定义 3.2.1, 它在点 $x \in S$ 处的次微分是

$$\begin{aligned} \partial \delta_S(x) &= \{x^* \in \mathcal{X}^* \mid \langle x^*, y - x \rangle \\ &\leq \delta_S(y) - \delta_S(x) \quad \forall y \in \mathcal{X}\}, \end{aligned}$$

即

$$\partial \delta_S(x) = \{x^* \in \mathcal{X}^* \mid \langle x^*, y - x \rangle \leq 0 \quad \forall y \in S\}. \quad (3.2.10)$$

由注 3.2.2 知, 对于任何 $x \in S$, (3.2.10) 右端的集合都是闭凸的. 此外, 对任意的 $\lambda > 0$, 显然有 $\langle \lambda x^*, y - x \rangle \leq 0 (\forall y \in S)$, 因而从 (3.2.10) 由定义 1.4.1 和定义 1.4.2 还得知 $\partial \delta_S(x)$ 是 \mathcal{X}^* 中的闭凸锥. 特别是, 当 $S = K$ 是凸锥时, 由 (3.2.10) 有

$$\partial \delta_S(\theta) = \{x^* \in \mathcal{X}^* \mid \langle x^*, y \rangle \leq 0 \quad \forall y \in K\}.$$

按定义 1.4.6 中的 (3), 它是凸锥 K 的负对偶锥, 即

$$K^\circ = \{x^* \in \mathcal{X}^* \mid \langle x^*, y \rangle \leq 0 \quad \forall y \in K\} = \partial \delta_S(\theta). \quad (3.2.11)$$

利用指示函数的次微分, 我们给出凸集的法锥和切锥概念如下.

定义 3.2.2 设 $S \subset \mathcal{X}$ 是非空凸集, 点 $x^0 \in S$.

(1) 集合

$$N_S(x^0) = \{x^* \in \mathcal{X}^* \mid \langle x^*, y - x^0 \rangle \leq 0 \quad \forall y \in S\} \quad (3.2.12)$$

称为凸集 S 在点 x^0 处的法锥, $x^* \in N_S(x^0)$ 称为 S 在点 x^0 处的法向量.

(2) 集合

$$T_S(x^0) = \{d \in \mathcal{H} \mid \langle x^*, d \rangle \leq 0 \quad \forall x^* \in N_S(x^0)\} \quad (3.2.13)$$

称为凸集 S 在点 x^0 处的切锥, $d \in T_S(x^0)$ 称为 S 在点 x^0 处的切向量.

注 3.2.4 设 $x \in S$, 由定义 3.2.2 不难验证 $N_S(x)$ 是 \mathcal{H}^* 中的锥. 对比 (3.2.12) 和 (3.2.10) 可知, 凸集 S 在点 x 处的法锥即是 S 上的指示函数在该点处的次微分: $N_S(x) = \partial \delta_S(x)$. 根据注 3.2.2, $\partial \delta_S(x)$ 是闭凸的, 故 $N_S(x)$ 是闭凸锥. 由 (3.2.13) 和 (3.2.11) 又得知, 凸集 S 在点 x 处的切锥即是该点处的法锥的负对偶锥: $T_S(x) = N_S(x)^-$.

注 3.2.5 在定义 3.2.2 中当 $\mathcal{H} = R^n$ 并把对偶积 $\langle \cdot, \cdot \rangle$ 理解为数量积时, 由 (3.2.12) 可知 $N_S(x^0)$ 就是与所有向量 $y - x^0$ ($y \in S$) 的夹角不小于 $\pi/2$ 的向量全体, 而 $T_S(x^0)$ 则是把原点移到 x^0 处后由 S 所生成的闭凸锥 (图 3.2.3). 若 S 在点 x^0 处是“光滑”的, 则 $T_S(x^0)$ 是过原点的半空间.

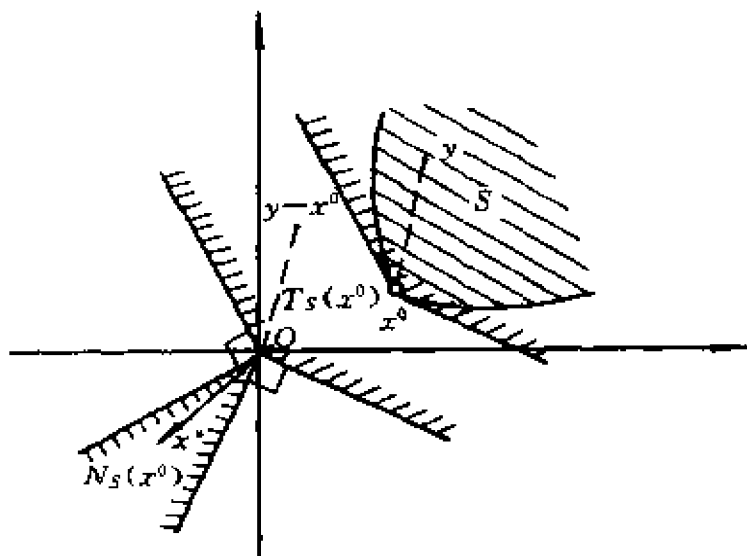


图 3.2.3

最后,介绍凸函数的次微分与共轭函数之间的关系.

定理 3.2.6 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $f^*: \mathcal{X}^* \rightarrow R \cup \{\pm\infty\}$ 是 f 的共轭函数, $x \in \text{dom} f$, 则 $x^* \in \partial f(x)$ 当且仅当

$$\langle x^*, x \rangle = f(x) + f^*(x^*). \quad (3.2.14)$$

证明 由定义 3.2.1, $x^* \in \partial f(x)$ 意味着

$$\langle x^*, y - x \rangle \leq f(y) - f(x) \quad \forall y \in \mathcal{X},$$

即

$$\langle x^*, x \rangle - f(x) \geq \langle x^*, y \rangle - f(y) \quad \forall y \in \mathcal{X}.$$

再由定义 2.2.4, 上式等价于

$$\langle x^*, x \rangle - f(x) = \sup_{y \in \mathcal{X}} \{ \langle x^*, y \rangle - f(y) \} = f^*(x^*),$$

此即 (3.2.14). \square

推论 3.2.7 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, f^* 和 f^{**} 分别是 f 的共轭函数和二次共轭函数, $x \in \text{dom} f$, 并且 $f(x) = f^{**}(x)$.

(1) $x^* \in \partial f(x)$ 当且仅当 $x \in \partial f^*(x^*)$.

(2) $\partial f(x) = \partial f^{**}(x)$.

证明 (1) 由定理 3.2.6 和 $f(x) = f^{**}(x)$, 则 $x^* \in \partial f(x)$ 等价于

$$\langle x^*, x \rangle = f^{**}(x) + f^*(x^*) = f^*(x^*) + (f^*)^*(x).$$

再利用定理 3.2.6 得知, 它又等价于 $x \in \partial f^*(x^*)$.

(2) 由 $f(x) = f^{**}(x)$ 知 $f(x^*) = (f^{**})^*(x^*)$. 于是, 由定理 3.2.6 得知 $x^* \in \partial f(x)$ 等价于

$$\langle x^*, x \rangle = f(x) + f^*(x^*) = f^{**}(x) + (f^{**})^*(x^*).$$

再由定理 3.2.6 即知, 它又等价于 $x^* \in \partial f^{**}(x)$. \square

推论 3.2.8 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, f^{**} 是

f 的二次共轭函数, $x \in \operatorname{dom} f$, 并且 $\partial f(x) \neq \emptyset$.

$$(1) f(x) = f^{**}(x),$$

$$(2) f(x) = \operatorname{cl} f(x).$$

证明 (1) 由 $\partial f(x) \neq \emptyset$, 设 $x^* \in \partial f(x)$, 据定理 3.2.6 有

$$\langle x^*, x \rangle = f(x) + f^*(x^*).$$

另外, 对 f^* 利用 (2.2.10) (Young-Fenchel 不等式), 则有

$$\langle x^*, x \rangle \leq f^{**}(x) + f^*(x^*).$$

由以上两式得到 $f(x) \leq f^{**}(x)$. 因为根据定理 2.2.12 的 (1) 知 f^{**} 是 f 的弱函数, 即有 $f^{**}(x) \leq f(x)$, 所以 $f(x) = f^{**}(x)$.

(2) 因 f^{**} 是 f 的弱函数, 故也有 $f^{**}(x) \leq \operatorname{cl} f(x) \leq f(x)$. 于是, 由 (1) 即知 $f(x) = \operatorname{cl} f(x)$. \square

§ 3.3 次微分的性质

设 \mathcal{X} 是线性拓扑空间. 由定义 3.2.1 我们知道, 凸函数 f 在点 $x \in \mathcal{X}$ 处的次微分 $\partial f(x)$ 是对偶空间 \mathcal{X}^* 中的一个集合, 因而它确定一个从 \mathcal{X} 到 \mathcal{X}^* 中集合的集值映射 $\partial f: \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$, $x \mapsto \partial f(x)$. 我们称此集值映射为 f 的次微分映射, 也简称次微分. 本节阐述次微分映射的运算性质、连续性和单调性.

先介绍次微分的简单运算性质.

定理 3.3.1 设 $f, f_1, f_2: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $x \in \mathcal{X}$.

$$(1) \text{ 若 } \alpha \geq 0, \text{ 则 } \partial(\alpha f)(x) = \alpha \partial f(x).$$

$$(2) \partial f_1(x) + \partial f_2(x) \subset \partial(f_1 + f_2)(x).$$

证明 (1) 当 $\alpha = 0$ 时, 是显然的. 设 $\alpha > 0$, 由定义 3.2.1 得知 $x^* \in \partial(\alpha f)(x)$ 意味着对任意的 $y \in \mathcal{X}$ 有

$$\langle x^*, y - x \rangle \leq \alpha f(y) - \alpha f(x),$$

或

$$\left\langle \frac{x^*}{\alpha}, y - x \right\rangle \leq f(y) - f(x).$$

再由定义 3.2.1 有 $\frac{x^*}{\alpha} \in \partial f(x)$, 即 $x^* \in \alpha \partial f(x)$.

(2) 设 $x^* \in \partial f_1(x) + \partial f_2(x)$, 则存在 x_1^* 和 x_2^* 使 $x^* = x_1^* + x_2^*$, 并且 $x_1^* \in \partial f_1(x)$ 和 $x_2^* \in \partial f_2(x)$. 由定义 3.2.1 可知

$$\langle x_1^*, y - x \rangle \leq f_1(y) - f_1(x) \quad \forall y \in \mathcal{H},$$

$$\langle x_2^*, y - x \rangle \leq f_2(y) - f_2(x) \quad \forall y \in \mathcal{H}.$$

将以上两式的两端相加, 得

$$\begin{aligned} \langle x_1^* + x_2^*, y - x \rangle &\leq [f_1(y) + f_2(y)] \\ &\quad - [f_1(x) + f_2(x)] \quad \forall y \in \mathcal{H}, \end{aligned}$$

再按定义 3.2.1, 于是得 $x^* = x_1^* + x_2^* \in \partial(f_1 + f_2)(x)$. \square

定理 3.3.2 设 $f: \mathcal{H} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $a \in \mathcal{H}^*$, $\beta \in R$, $l(x) = \langle a, x \rangle + \beta$ 是 \mathcal{H} 上的仿射函数, 则

$$\partial(f + l)(x) = \partial f(x) + \partial l(x) \quad \forall x \in \mathcal{H}.$$

证明 由定义 3.2.1 可知, $x^* \in \partial(f + l)(x)$ 意味着

$$\begin{aligned} \langle x^*, y - x \rangle &\leq [f(y) + \langle a, y \rangle + \beta] \\ &\quad - [f(x) + \langle a, x \rangle + \beta] \quad \forall y \in \mathcal{H}, \end{aligned}$$

或即

$$\langle x^* - a, y - x \rangle \leq f(y) - f(x) \quad \forall y \in \mathcal{H}.$$

于是, 再按定义 3.2.1 有 $x^* - a \in \partial f(x)$, 由例 3.2.2 知 $\partial l(x) = \{a\}$, 因而得 $x^* \in \partial f(x) + \partial l(x)$. \square

定理 3.3.3 设 $f_i: \mathcal{H} \rightarrow R \cup \{+\infty\} (i = 1, \dots, m)$ 是正常凸函数,

$$f(x) = \max_{1 \leq i \leq m} f_i(x), \quad x \in \mathcal{H},$$

$I(x) = \{i \in \{1, \dots, m\} | f_i(x) = f(x), x \in \mathcal{X}\}$. 若各 f_i 在点 $x^0 \in \text{dom} f$ 处是连续的, 则

$$\partial f(x^0) = \text{co} \bigcup_{i \in I(x^0)} \partial f_i(x^0).$$

证明 因为各 $f_i (i = 1, \dots, m)$ 在 x^0 处是连续的, 所以由定理 3.2.2 有

$$\begin{aligned} \partial f_i(x^0) &= \{x^* \in \mathcal{X}^* | \langle x^*, d \rangle \leq f'_{i+}(x^0; d) \quad \forall d \in \mathcal{X}\}, \\ i &= 1, \dots, m. \end{aligned} \quad (3.3.1)$$

又由 f_i 在 x^0 处是连续的得知, 在 x^0 的某邻域内有

$$f(x) = \max_{i \in I(x^0)} f_i(x), \quad (3.3.2)$$

并且它在该点处是连续的. 于是同样由定理 3.2.2 又有

$$\begin{aligned} \partial f(x^0) &= \{x^* \in \mathcal{X}^* | \langle x^*, d \rangle \leq f'_+(x^0; d) \quad \forall d \in \mathcal{X}\}. \end{aligned} \quad (3.3.3)$$

现在证明

$$f'_+(x^0; d) = \max_{i \in I(x^0)} f'_{i+}(x^0; d), \quad d \in \mathcal{X}. \quad (3.3.4)$$

事实上, 当 $i \in I(x^0)$ 时, 由定义 3.1.1 的(1)和(3.3.2)有

$$\begin{aligned} f'_{i+}(x^0; d) &= \lim_{t \rightarrow 0^+} \frac{f_i(x^0 + td) - f_i(x^0)}{t} \\ &\leq \lim_{t \rightarrow 0^+} \frac{f(x^0 + td) - f(x^0)}{t} \\ &= f'_+(x^0; d), \end{aligned}$$

因此

$$f'_-(x^0; d) \geq \max_{i \in I(x^0)} f'_{i+}(x^0; d). \quad (3.3.5)$$

另一方面, 从定义 3.1.1 的(1)可知有正数列 $\{t_k\}$, $t_k \rightarrow 0 (k \rightarrow$

∞), 并且

$$f'_+(x^0; d) = \lim_{k \rightarrow \infty} \frac{f(x^0 + t_k d) - f(x^0)}{t_k}. \quad (3.3.6)$$

由(3.3.2)得知, 至少存在一个 $\bar{i} \in I(x^0)$ 使对无限多个 k 有

$$f(x^0 + t_k d) = f_{\bar{i}}(x^0 + t_k d).$$

由于对这个 $\bar{i} \in I(x^0)$ 根据(3.3.6)有

$$f'_+(x^0; d) = \lim_{k \rightarrow \infty} \frac{f_{\bar{i}}(x^0 + t_k d) - f_{\bar{i}}(x^0)}{t_k} = f'_{\bar{i}+}(x^0; d),$$

故由(3.3.5)知(3.3.4)成立.

根据(3.3.4)和注 3.2.3 可知, 对 $d \in \mathcal{H}$ 有

$$\begin{aligned} f'_+(x^0; d) &= \max_{i \in I(x^0)} f'_{i+}(x^0; d) \\ &= \max_{i \in I(x^0)} \sup_{x^* \in \partial f_i(x^0)} \langle x^*, d \rangle, \end{aligned}$$

即

$$f'_+(x^0; d) = \sup_{\substack{x^* \in \bigcup_{i \in I(x^0)} \partial f_i(x^0)}} \langle x^*, d \rangle, \quad d \in \mathcal{H}.$$

由此, 按定义 1.3.4 得知 $f'_+(x^0; d)$ 关于 d 是集合 $\bigcup_{i \in I(x^0)} \partial f_i(x^0)$ 的

支撑函数. 因为显然有 $\bigcup_{i \in I(x^0)} \partial f_i(x^0) \neq \emptyset$, 所以由定理 1.3.13

和(3.3.3)得到

$$\text{co } \bigcup_{i \in I(x^0)} \partial f_i(x^0) = \{x^* \in \mathcal{H}^* \mid \langle x^*, d \rangle \leq f'_+(x^0; d)$$

$$\forall d \in \mathcal{H}\} = \partial f(x^0).$$

最后, 证明 $\text{co } \bigcup_{i \in I(x^0)} \partial f_i(x^0)$ 是闭集, 从上式即可得定理的结

论. 为此, 作函数 $g: R^m \times (\mathcal{H}^*)^m \rightarrow R$,

$$g(\lambda, x_1^*, \dots, x_m^*) = \sum_{i=1}^m \lambda_i x_i^*, \quad (3.3.7)$$

其中 $\lambda = (\lambda_1, \dots, \lambda_m)^T \in R^m$. 记

$$\Lambda = \left\{ \lambda = (\lambda_1, \dots, \lambda_m)^T \in R^m \mid \lambda_i \geq 0 \right. \\ \left. (i = 1, \dots, m), \sum_{i=1}^m \lambda_i = 1 \right\},$$

由(3.3.7)可以推知

$$g\left(\Lambda \times \left\{ \prod_{i \in I(x^0)} \partial f_i(x^0) \right\}\right) = \text{co} \bigcup_{i \in I(x^0)} \partial f_i(x^0). \quad (3.3.8)$$

因为 Λ 和 $\partial f_i(x^0) (i \in I(x^0))$ 都是闭集, 并且由(3.3.7)知 g 关于 λ 和 $x_i^* (i = 1, \dots, m)$ 是连续的, 所以根据(3.3.8)即得 $\text{co} \bigcup_{i \in I(x^0)} \partial f_i(x^0)$ 是闭集. \square

下面是次微分运算的基本定理.

定理 3.3.4 设 $f_1, f_2: \mathcal{H} \rightarrow R \cup \{+\infty\}$ 是正常凸函数. 若存在点 $x^0 \in (\text{dom} f_1) \cap (\text{dom} f_2)$, 使得 f_1 在点 x^0 处是连续的, 则

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall x \in \mathcal{H}. \quad (3.3.9)$$

证明 由定理 3.3.1 的(2), 只需再证明

$$\partial(f_1 + f_2)(x) \subset \partial f_1(x) + \partial f_2(x).$$

为此, 设 $x^* \in \partial(f_1 + f_2)(x)$, 我们证明有 $x^* \in \partial f_1(x) + \partial f_2(x)$. 不妨设 $x^* = \theta, x = \theta, f_1(\theta) + f_2(\theta) = 0$, 因为否则, 可令

$$g_1(z) = f_1(x + z) - f_1(x) - \langle x^*, z \rangle,$$

$$g_2(z) = f_2(x + z) - f_2(x),$$

将问题转为对 g_1 和 g_2 来讨论.

以下设 $\theta \in \partial(f_1 + f_2)(\theta)$, 我们来证明 $\theta \in \partial f_1(\theta) + \partial f_2(\theta)$. 从 $\theta \in \partial(f_1 + f_2)(\theta)$, 由定义 3.2.1 知

$$\langle \theta, y - \theta \rangle \leq (f_1 + f_2)(y) - (f_1 + f_2)(\theta) \quad \forall y \in \mathcal{X},$$

从而有

$$\inf_{y \in \mathcal{X}} [f_1(y) + f_2(y)] = f_1(\theta) + f_2(\theta) = 0. \quad (3.3.10)$$

作集合

$$S_1 = \text{epi} f_1 = \{(x, \eta) \in \mathcal{X} \times R \mid \eta \geq f_1(x)\}, \quad (3.3.11)$$

$$S_2 = \{(x, \eta) \in \mathcal{X} \times R \mid -f_2(x) \geq \eta\}. \quad (3.3.12)$$

由 $(\text{dom} f_1) \cap (\text{dom} f_2) \neq \emptyset$ 和 f_1 是凸函数, 根据定理 2.1.5 知 S_1 是凸集. 同理, 由 f_2 是凸函数可推知 S_2 是非空凸集. 又由 f_1 在点 x^0 处是连续的, 根据定理 2.2.1, 它在 $\text{int}(\text{dom} f_1)$ 内是连续的, 于是由推论 2.2.2 知 $\text{int} S_1 = \text{int}(\text{epi} f_1) \neq \emptyset$. 此外, 可以证明有 $\text{int} S_1 \cap S_2 = \emptyset$. 事实上, 否则, 若存在 $(\bar{x}, \bar{\eta}) \in \text{int} S_1 \cap S_2$, 则由 (3.3.11) 和 (3.3.12) 有

$$f_1(\bar{x}) < \bar{\eta} \leq -f_2(\bar{x}).$$

根据此, 导致

$$\inf_{y \in \mathcal{X}} [f_1(y) + f_2(y)] \leq f_1(\bar{x}) + f_2(\bar{x}) < 0,$$

它与 (3.3.10) 相矛盾. 根据以上得到的 $\text{int} S_1 \neq \emptyset$, $S_2 \neq \emptyset$ 和 $\text{int} S_1 \cap S_2 = \emptyset$, 利用定理 1.3.7 的 (2), 可知存在闭超平面严格分离 $\text{int} S_1$ 和 S_2 . 按定义 1.3.2 的 (2), 也即存在 $(x^*, \eta^*) \in (\mathcal{X}^* \times R) \setminus \{\theta\}$ 有

$$\langle (x^*, \eta^*), (x^1, \eta_1) \rangle < \langle (x^*, \eta^*), (x^2, \eta_2) \rangle$$

$$\forall (x^1, \eta_1) \in \text{int} S_1, \forall (x^2, \eta_2) \in S_2,$$

或

$$\langle \mathbf{x}^*, \mathbf{x}^1 \rangle + \eta^* \eta_1 < \langle \mathbf{x}^*, \mathbf{x}^2 \rangle + \eta^* \eta_2$$

$$\forall (\mathbf{x}^1, \eta_1) \in \text{int} S_1, \quad \forall (\mathbf{x}^2, \eta_2) \in S_2. \quad (3.3.13)$$

(3.3.13)左端的 η_1 可以任意大, 故 $\eta^* \leq 0$. 但若有 $\eta^* = 0$, 则从 $\mathbf{x}^0 \in (\text{dom} f_1) \cap (\text{dom} f_2)$ 且 f_1 在点 \mathbf{x}^0 处是连续的可知, 存在 $\varepsilon > 0$ 有

$$(\mathbf{x}^0, f_1(\mathbf{x}^0) + \varepsilon) \in \text{int} S_1, \quad (\mathbf{x}^0, -f_2(\mathbf{x}^0)) \in S_2,$$

于是由(3.3.13)导致矛盾:

$$\langle \mathbf{x}^*, \mathbf{x}^0 \rangle < \langle \mathbf{x}^*, \mathbf{x}^0 \rangle.$$

因此, 我们得知 $\eta^* < 0$.

令 $\mathbf{x}_1^* = -\frac{\mathbf{x}^*}{\eta^*}$, 由(3.3.13)注意到 $\eta^* < 0$, 得

$$-\langle \mathbf{x}_1^*, \mathbf{x}^1 \rangle + \eta_1 \geq -\langle \mathbf{x}_1^*, \mathbf{x}^2 \rangle + \eta_2$$

$$\forall (\mathbf{x}^1, \eta_1) \in S_1, \quad \forall (\mathbf{x}^2, \eta_2) \in S_2.$$

特别是, 取 $\eta_1 = f_1(\mathbf{x}^1)$, $\eta_2 = -f_2(\mathbf{x}^2)$, 则有

$$-\langle \mathbf{x}_1^*, \mathbf{x}^1 \rangle + f_1(\mathbf{x}^1) \geq -\langle \mathbf{x}_1^*, \mathbf{x}^2 \rangle - f_2(\mathbf{x}^2)$$

$$\forall \mathbf{x}^1 \in \text{dom} f_1, \quad \forall \mathbf{x}^2 \in \text{dom} f_2. \quad (3.3.14)$$

在(3.3.14)中取 $\mathbf{x}^2 = \mathbf{0}$, 因为 $f_1(\mathbf{0}) = f_2(\mathbf{0}) = 0$, 故得

$$\langle \mathbf{x}_1^*, \mathbf{x}^1 \rangle \leq f_1(\mathbf{x}^1) = f_1(\mathbf{x}^1) - f_1(\mathbf{0}) \quad \forall \mathbf{x}^1 \in \text{dom} f_1.$$

据此, 由定义 3.2.1 即知 $\mathbf{x}_1^* \in \partial f_1(\mathbf{0})$. 在(3.3.14)中再取 $\mathbf{x}^1 = \mathbf{0}$, 同理有

$$-\langle \mathbf{x}_1^*, \mathbf{x}^2 \rangle \leq f_2(\mathbf{x}^2) = f_2(\mathbf{x}^2) - f_2(\mathbf{0}) \quad \forall \mathbf{x}^2 \in \text{dom} f_2,$$

于是得 $-\mathbf{x}_1^* \in \partial f_2(\mathbf{0})$. 因此得到 $\mathbf{0} = \mathbf{x}_1^* - \mathbf{x}_1^* \in \partial f_1(\mathbf{0}) + \partial f_2(\mathbf{0})$. \square

以下讨论次微分映射的连续性. 为此, 先给出集值映射连续性

的定义.

设 \mathscr{U} 是线性拓扑空间.

定义 3.3.1 设点 $x \in \mathscr{X}$, 集合 $\phi(x) \subset \mathscr{U}$, $\phi: \mathscr{X} \rightarrow 2^{\mathscr{U}}$, $x \mapsto \phi(x)$ 是集值映射.

(1) 若点 $x^0 \in \mathscr{X}$, 并且对任意的点列 $\{x^k\} \subset \mathscr{X}$, $x^k \rightarrow x^0$ ($k \rightarrow \infty$), 以及 $y^k \in \phi(x^k)$, $y^k \rightarrow y^0$, 有 $y^0 \in \phi(x^0)$, 则称集值映射 ϕ 在点 $x^0 \in \mathscr{X}$ 处是上半连续的. 若 ϕ 在集合 $S \subset \mathscr{X}$ 的每一点处都是上半连续的, 则称 ϕ 是 S 上的上半连续集值映射, 或 ϕ 在 S 上是上半连续的.

(2) 若点 $x^0 \in \mathscr{X}$, 并且对任意的点列 $\{x^k\} \subset \mathscr{X}$, $x^k \rightarrow x^0$ ($k \rightarrow \infty$), 以及 $y^0 \in \phi(x^0)$, 存在正整数 M 和点列 $\{y^k\} \subset \mathscr{U}$ 有 $y^k \in \phi(x^k)$ ($k \geq M$) 和 $y^k \rightarrow y^0$, 则称集值映射 ϕ 在点 $x^0 \in \mathscr{X}$ 处是下半连续的. 若 ϕ 在集合 $S \subset \mathscr{X}$ 的每一点处都是下半连续的, 则称 ϕ 是 S 上的下半连续集值映射, 或 ϕ 在 S 上是下半连续的.

(3) 若 ϕ 在点 $x^0 \in \mathscr{X}$ 处既是上半连续的又是下半连续的, 则称集值映射 ϕ 在点 x^0 处是连续的. 若 ϕ 在集合 $S \subset \mathscr{X}$ 的每一点处都是连续的, 则称 ϕ 是 S 上的连续集值映射, 或 ϕ 在 S 上是连续的.

注 3.3.1 当 $\mathscr{U} = R$, 并且 $\phi = f: \mathscr{X} \rightarrow R$ 是实值函数时, 定义 3.3.1 的 (1) 意即对任意的 $\{x^k\} \subset \mathscr{X}$, $x^k \rightarrow x^0$ ($k \rightarrow \infty$) 有 $f(x^k) \rightarrow f(x^0)$, 也即 $\lim_{x^k \rightarrow x^0} f(x^k) = f(x^0)$. 由 $\{x^k\}$ 的任意性, 可知它等价于对任意的 $\varepsilon > 0$ 存在 x^0 的邻域 $U(x^0) \subset \mathscr{X}$, 有

$$|f(x) - f(x^0)| < \varepsilon \quad \forall x \in U(x^0).$$

对照定义 2.2.1, 即 f 在点 x^0 处是连续的. 类似地, 定义 3.3.1 的 (2) 意即对任意的 $\{x^k\} \subset \mathscr{X}$, $x^k \rightarrow x^0$ ($k \rightarrow \infty$), 存在 M 和数列 $\{f(x^k)\} \subset R$, 当 $k \geq M$ 时, 有 $f(x^k) \rightarrow f(x^0)$, 即 $\lim_{x^k \rightarrow x^0} f(x^k) = f(x^0)$. 同理, 它也即 f 在点 x^0 处是连续的. 因此, 当集值映射退化为实值函数时, 它的上半连续性和下半连续性 (这时等同) 以及连

续性都是实值函数(实泛函)的连续性. 但应指出, 对于线性拓扑空间上的一般集值映射而言, 它的上半连续性和下半连续性是互不包含的.

定理 3.3.5 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数. 若 f 在点 $\mathbf{x}^0 \in \text{dom} f$ 处是连续的, 则 ∂f 在 \mathbf{x}^0 处是上半连续的.

证明 从 f 在点 \mathbf{x}^0 处是连续的, 由定义 2.2.1 知 f 在 \mathbf{x}^0 的某邻域 $U_1(\mathbf{x}^0)$ 内上有界. 因此, 由定理 2.2.1, 它在 $U_1(\mathbf{x}^0)$ 内是连续的. 对于 $\mathbf{x} \in U_1(\mathbf{x}^0)$, 由定义 3.1.1 的(1)有

$$f'_+(\mathbf{x}; \mathbf{d}) = \lim_{t \rightarrow 0^+} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t} \quad \forall \mathbf{d} \in \mathcal{X}, \quad (3.3.15)$$

故对任何 $\mathbf{d} \in \mathcal{X}$, 函数

$$F(\mathbf{x}, t) = \begin{cases} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}, & t > 0; \\ f'_+(\mathbf{x}; \mathbf{d}), & t = 0, \end{cases}$$

在 $U_1(\mathbf{x}^0) \times [0, \infty)$ 内是连续的. 据此, 按定义 2.2.1 可知, 对任意的 $\varepsilon > 0$ 存在 \mathbf{x}^0 的邻域 $U(\mathbf{x}^0) \subset U_1(\mathbf{x}^0)$ 和 $\delta_1 > 0$, 使得

$$\left| \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t} - f'_+(\mathbf{x}^0; \mathbf{d}) \right| < \frac{\varepsilon}{2} \quad \forall \mathbf{x} \in U(\mathbf{x}^0), \forall t \in (0, \delta_1). \quad (3.3.16)$$

另外, 从(3.3.15)得知对上述的 $\varepsilon > 0$ 又存在 $\delta_2 > 0$, 有

$$\left| \frac{f(\mathbf{x}^0 + t\mathbf{d}) - f(\mathbf{x}^0)}{t} - f'_+(\mathbf{x}^0; \mathbf{d}) \right| < \frac{\varepsilon}{2} \quad \forall t \in (0, \delta_2). \quad (3.3.17)$$

取 $\delta = \min\{\delta_1, \delta_2\}$, 由(3.3.16)和(3.3.17)可得对任意的 $\mathbf{x} \in U(\mathbf{x}^0)$ 和任意的 $t \in (0, \delta)$, 有

$$\frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t} < f'_+(\mathbf{x}^0; \mathbf{d}) + \frac{\varepsilon}{2}$$

$$< \frac{f(\mathbf{x}^0 + t\mathbf{d}) - f(\mathbf{x}^0)}{t} + \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

即

$$\frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t} < \frac{f(\mathbf{x}^0 + t\mathbf{d}) - f(\mathbf{x}^0)}{t} + \varepsilon$$

$$\forall \mathbf{x} \in U(\mathbf{x}^0), \forall t \in (0, \delta).$$

由此, 根据定义 3.1.1 的(1), 对任意的 $\mathbf{x} \in U(\mathbf{x}^0)$ 有

$$\begin{aligned} f'_+(\mathbf{x}; \mathbf{d}) &= \lim_{t \rightarrow 0^+} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t} \\ &\leq \lim_{t \rightarrow 0^+} \frac{f(\mathbf{x}^0 + t\mathbf{d}) - f(\mathbf{x}^0)}{t} + \varepsilon \\ &= f'_+(\mathbf{x}^0; \mathbf{d}) + \varepsilon, \end{aligned}$$

从而

$$\inf_{U(\mathbf{x}^0)} \sup_{\mathbf{x} \in U(\mathbf{x}^0)} f'_+(\mathbf{x}; \mathbf{d}) \leq f'_+(\mathbf{x}^0; \mathbf{d}) + \varepsilon.$$

令 $\varepsilon \rightarrow 0$, 得到

$$\inf_{U(\mathbf{x}^0)} \sup_{\mathbf{x} \in U(\mathbf{x}^0)} f'_-(\mathbf{x}; \mathbf{d}) \leq f'_-(\mathbf{x}^0; \mathbf{d}). \quad (3.3.18)$$

根据注 2.1.3, 即知 $f'_-(\mathbf{x}; \mathbf{d})$ 关于 \mathbf{x} 在 \mathbf{x}^0 处是上半连续的.

因为 f 是正常凸函数并且在 $U(\mathbf{x}^0)$ 内是连续的, 由定理 3.2.1 可知, 对任意的 $\mathbf{x} \in U(\mathbf{x}^0)$ 有 $\partial f(\mathbf{x}) \neq \emptyset$. 设 $\mathbf{x}^* \in \partial f(\mathbf{x})$, 根据定理 3.2.2 得到

$$\langle \mathbf{x}^*, \mathbf{d} \rangle \leq f'_-(\mathbf{x}; \mathbf{d}) \quad \forall \mathbf{d} \in \mathcal{H}, \forall \mathbf{x} \in U(\mathbf{x}^0). \quad (3.3.19)$$

以下证明集值映射 ∂f 在 \mathbf{x}^0 处是上半连续的. 为此, 任取点列 $\{\mathbf{x}_k\} \subset \mathcal{H}$, $\mathbf{x}_k \rightarrow \mathbf{x}^0$ ($k \rightarrow \infty$), 以及 $\mathbf{x}_k^* \in \partial f(\mathbf{x}_k)$, $\mathbf{x}_k^* \rightarrow \mathbf{x}^*$. 从 (3.3.19) 有

$$\langle \mathbf{x}_k^*, \mathbf{d} \rangle \leq f'_-(\mathbf{x}_k; \mathbf{d}) \quad \forall \mathbf{d} \in \mathcal{H}, \forall \mathbf{x}_k \in U(\mathbf{x}^0).$$

据此,由(3.3.18)知,对任意的 $d \in \mathcal{H}$ 有

$$\begin{aligned}\langle x^*, d \rangle &= \lim_{k \rightarrow \infty} \langle x_k^*, d \rangle \leq \inf_{U(x^0)} \sup_{x_k \in U(x^0)} f'_-(x_k; d) \\ &\leq \inf_{U(x^0)} \sup_{x \in U(x^0)} f'_+(x; d) \\ &\leq f'_+(x^0; d),\end{aligned}$$

于是根据定理 3.2.2 可知 $x^* \in \partial f(x^0)$. 最后,由定义 3.3.1 的 (1), 使得 ∂f 在 x^0 处是上半连续的. \square

最后,阐述次微分映射的单调性.

定义 3.3.2 设点 $x \in \mathcal{X}$, 集合 $\phi(x) \subset \mathcal{X}^*$, $\phi: \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$, $x \mapsto \phi(x)$ 是集值映射. 若对任意的 $x, y \in \mathcal{X}$ 有

$$\langle y^* - x^*, y - x \rangle \geq 0 \quad \forall x^* \in \phi(x), \forall y^* \in \phi(y), \quad (3.3.20)$$

则称 ϕ 是 \mathcal{X} 上的单调集值映射, 或称集值映射 ϕ 在 \mathcal{X} 上是单调的.

注 3.3.2 当 $\mathcal{X} = R$ 和 $2^{\mathcal{X}} = R$, ϕ 是单变量函数 $\phi = f: R \rightarrow R$ 时, 定义 3.3.2 中的 $x^* = f(x)$, $y^* = f(y)$, (3.3.20) 即为

$$\langle f(y) - f(x), y - x \rangle \geq 0.$$

由此, 当 $x \leq y$ 时有 $f(x) \leq f(y)$, 这也即 f 是 R 上的单调函数.

定理 3.3.6 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是连续的正常凸函数, 则 ∂f 在 \mathcal{X} 上是单调的.

证明 对于任意的 $x, y \in \mathcal{X}$, 任取 $x^* \in \partial f(x)$ 和 $y^* \in \partial f(y)$, 由定义 3.2.1 有

$$\langle x^*, y - x \rangle \leq f(y) - f(x) \quad \forall y \in \mathcal{X} \quad (3.3.21)$$

和

$$\langle y^*, x - y \rangle \leq f(x) - f(y) \quad \forall x \in \mathcal{X}. \quad (3.3.22)$$

由(3.3.22)有 $\langle -y^*, y - x \rangle \leq f(x) - f(y)$. 将它与(3.3.21)的两端相加得 $\langle x^* - y^*, y - x \rangle \leq 0$, 注意, $x^* \in \partial f(x)$ 和 $y^* \in$

$\partial f(y)$ 是任取的, 有

$$\langle x^* - y^*, y - x \rangle \geq 0 \quad \forall x^* \in \partial f(x), \forall y^* \in \partial f(y).$$

因此, 由定义 3.3.2 即知 ∂f 在 \mathcal{X} 上是单调的. \square

进而, 下面考虑更强的单调性.

定义 3.3.3 设点 $x \in \mathcal{X}$, 集合 $\phi(x) \subset \mathcal{X}^*$, $\phi: \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$, $x \mapsto \phi(x)$ 是集值映射. 若对任何正整数 $m \geq 2$ 和任意的 $x^i \in \mathcal{X}$ ($i = 1, \dots, m$) 有

$$\begin{aligned} & \langle x_1^*, x^2 - x^1 \rangle + \langle x_2^*, x^3 - x^2 \rangle + \dots + \langle x_m^*, x^1 - x^m \rangle \\ & \leq 0 \quad \forall x_1^* \in \phi(x^1), \dots, \forall x_m^* \in \phi(x^m), \end{aligned} \quad (3.3.23)$$

则称 ϕ 是 \mathcal{X} 上的循环单调集值映射, 或集值映射 ϕ 在 \mathcal{X} 上是循环单调的.

显然, 在定义 3.3.3 中当 $m = 2$ 时, (3.3.23) 即

$$\langle x_1^*, x^2 - x^1 \rangle + \langle x_2^*, x^1 - x^2 \rangle \leq 0$$

$$\forall x_1^* \in \phi(x^1), \forall x_2^* \in \phi(x^2),$$

或

$$\langle x_2^* - x_1^*, x^2 - x^1 \rangle \geq 0 \quad \forall x_1^* \in \phi(x^1), \forall x_2^* \in \phi(x^2).$$

于是, 由定义 3.3.2 可知, 这时 ϕ 在 \mathcal{X} 上是单调的.

定理 3.3.7 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是连续的正常凸函数, 则 ∂f 在 \mathcal{X} 上是循环单调的.

证明 对任意的 $x^i \in \mathcal{X}$ ($i = 1, \dots, m$), 任取 $x_i^* \in \partial f(x^i)$, 由定义 3.2.1 有

$$\langle x_1^*, x^2 - x^1 \rangle \leq f(x^2) - f(x^1),$$

$$\langle x_2^*, x^3 - x^2 \rangle \leq f(x^3) - f(x^2),$$

.....

$$\langle x_m^*, x^1 - x^m \rangle \leq f(x^1) - f(x^m).$$

将以上各式两端相加, 注意各 $\mathbf{x}_i' \in \partial f(\mathbf{x}')$ 是任取的, 即得 (3.3.23). 由此, 按定义 3.3.3 即可知 ∂f 在 \mathcal{K} 上是循环单调的. \square

§ 3.4 凸规划的最优性条件

利用凸函数的次微分, 本节论述凸规划问题的最优解要满足的最优性条件. 这是数学规划理论中最基本的问题.

设 \mathcal{V} 是线性空间, $f: \mathcal{V} \rightarrow R \cup \{\pm \infty\}$, $g_i: \mathcal{V} \rightarrow R \cup \{\pm \infty\}$ ($i = 1, \dots, p$) 和 $h_j: \mathcal{V} \rightarrow R \cup \{\pm \infty\}$ ($j = 1, \dots, q$) 都是广义实值函数, 则一般形式的数学规划问题记作

$$\begin{cases} \min f(\mathbf{x}) \\ \text{s. t. } g_i(\mathbf{x}) \leq 0, i = 1, \dots, p, \\ h_j(\mathbf{x}) = 0, j = 1, \dots, q. \end{cases} \quad (\text{MP})$$

上式中的 f 称为 (MP) 的目标函数, $\min f(\mathbf{x})$ 表示对 $f(\mathbf{x})$ 的极小化; g_i ($i = 1, \dots, p$) 和 h_j ($j = 1, \dots, q$) 称为约束函数, s. t. 表示受约束于, 而称

$$g_i(\mathbf{x}) \leq 0, i = 1, \dots, p$$

是 (MP) 的不等式约束条件;

$$h_j(\mathbf{x}) = 0, j = 1, \dots, q$$

是 (MP) 的等式约束条件. 记

$$X = \{\mathbf{x} \in \mathcal{V} \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, p; h_j(\mathbf{x}) = 0, j = 1, \dots, q\}, \quad (3.4.1)$$

则问题 (MP) 也可简记作

$$\min_{\mathbf{x} \in X} f(\mathbf{x}), \quad (\text{MP})$$

其中 $X \subset \mathcal{V}$ 称为(MP)的约束集.

特别是,当 $X = \mathcal{V}$ 时,称(MP)是无约束极小化问题,并记它为

$$\min f(x), \quad (\text{UMP})$$

当 $X \neq \emptyset$ 是 \mathcal{V} 的真子集时,则称(MP)是约束极小化问题.

若问题(MP)中的目标函数 f 是凸函数,并且约束集 X 是凸集,则称它是凸规划问题或凸规划.若(MP)中的目标函数 f 以及约束函数 $g_i (i = 1, \dots, p)$ 和 $h_j (j = 1, \dots, q)$ 都是线性函数,则称它是线性规划问题或线性规划.

注 3.4.1 若(MP)中的 $f, g_i (i = 1, \dots, p)$ 是凸函数, $h_j (j = 1, \dots, q)$ 是仿射函数,则(MP)是一凸规划.事实上,记

$$S_1 = \bigcap_{i=1}^p \{x \in \mathcal{V} \mid g_i(x) \leq 0\},$$

$$S_2 = \bigcap_{j=1}^q \{x \in \mathcal{V} \mid h_j(x) = 0\},$$

由(3.4.1)知

$$X = S_1 \cap S_2.$$

因为 $g_i (i = 1, \dots, p)$ 和 $h_j (j = 1, \dots, q)$ 都是凸函数,根据定理 2.1.2 知,它们相应的水平集 $\{x \in \mathcal{V} \mid g_i(x) \leq 0\} (i = 1, \dots, p)$ 和 $\{x \in \mathcal{V} \mid h_j(x) = 0\} (j = 1, \dots, q)$ 都是凸集,所以由定理 1.1.2 的(2)得知它们的交集 X 是凸集.此外,由于线性函数也是凸函数,故线性规划也是凸规划.

对于凸规划问题(MP),考虑它的约束集 X 上的指示函数

$$\delta_X(x) = \begin{cases} 0, & x \in X; \\ +\infty, & x \in \mathcal{V} \setminus X, \end{cases}$$

则约束极小化问题(MP)可以转化为等价的无约束极小化问题:

$$\min [f(x) + \delta_X(x)]. \quad (3.4.2)$$

定义 3.4.1 设集合 $X \subset \mathcal{X}$ 非空, $f: \mathcal{X} \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数. 若 $\bar{x} \in X$, 并且

$$f(\bar{x}) \leq f(x) \quad \forall x \in X,$$

则称 \bar{x} 是数学规划问题 (MP) 的最优解 (或极小解), $f(\bar{x})$ 是 (MP) 的最优值 (或极小值). (MP) 的所有最优解组成的集合称为 (MP) 的最优解集, 记作 \tilde{X} .

对于线性拓扑空间 \mathcal{X} 上的无约束凸规划问题 (UMP), 它的最优解有如下简单的最优性充分必要条件.

定理 3.4.1 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, 则 \bar{x} 是 (UMP) 的最优解当且仅当 $\mathbf{0} \in \partial f(\bar{x})$.

证明 由定义 3.4.1, 注意到这时 $X = \mathcal{X}$, 则 \bar{x} 是 (UMP) 的最优解意味着

$$f(x) - f(\bar{x}) \geq 0 \quad \forall x \in \mathcal{X},$$

即

$$\langle \mathbf{0}, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \quad \forall x \in \mathcal{X}.$$

依据定义 3.2.1, 它与 $\mathbf{0} \in \partial f(\bar{x})$ 等价. \square

推论 3.4.2 设 $f: R^n \rightarrow R$ 是正常凸函数. 若 f 在 R^n 上可微, 则 \bar{x} 是 (UMP) 的最优解当且仅当 $\nabla f(\bar{x}) = \mathbf{0}$.

证明 因为 f 是 R^n 上的可微函数, 由推论 3.2.5 可知, 对任意的 $x \in R^n$, 有 $\partial f(x) = \{\nabla f(x)\}$. 由此, 根据定理 3.4.1 即可得到 $\nabla f(\bar{x}) = \mathbf{0}$. \square

下面的定理指出, 凸规划问题的最优解集必是凸集.

定理 3.4.3 设 $X \subset \mathcal{X}$ 是非空凸集, $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, 则 (MP) 的最优解集 \tilde{X} 是凸集.

证明 设 $\tilde{X} \neq \emptyset$, 则有最优解 $\bar{x} \in \tilde{X}$. 记 (MP) 的最优值 $\bar{f} = f(\bar{x})$, 由定义 3.4.1 知

$$\tilde{X} = \{x \in X \mid f(x) \leq \bar{f}\}.$$

因为 X 是非空凸集, f 是凸函数, 所以由定理 2.1.2 得知, 对于 $\tilde{f} \in R$, 其相应的水平集 \tilde{X} 是凸集. \square

为论述约束凸规划问题 (MP) 的最优解要满足的最优性条件, 我们先给出如下的引理.

引理 3.4.4 设 $F: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 是正常凸函数, $h_j: \mathcal{X} \rightarrow R (j = 1, \dots, q)$ 是仿射函数, $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_q(\mathbf{x}))^T$, 并且 $\mathbf{0} \in (\mathbf{h}(\text{dom}F))^i$, 则

$$F(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \{\mathbf{x} \in \mathcal{X} | \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \quad (3.4.3)$$

当且仅当存在 $w > 0$ 和 $\mu_j \in R (j = 1, \dots, q)$ 使得

$$wF(\mathbf{x}) + \sum_{j=1}^q \mu_j h_j(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \text{dom}F. \quad (3.4.4)$$

证明 充分性. 由 $\mathbf{0} \in (\mathbf{h}(\text{dom}F))^i$ 可知存在 $\mathbf{x} \in \text{dom}F$ 使 $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, 据此, 从 (3.4.4) 我们有

$$wF(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \{\mathbf{x} \in \mathcal{X} | \mathbf{h}(\mathbf{x}) = \mathbf{0}\}.$$

因为 $w > 0$, 所以得到 (3.4.3).

必要性. 作向量函数 $G: \text{dom}F \rightarrow R \times R^q$,

$$G(\mathbf{x}) = (F(\mathbf{x}), \mathbf{h}(\mathbf{x})), \quad (3.4.5)$$

记

$$A = \{(y, \mathbf{z}) \in R \times R^q | y \geq 0, \mathbf{z} = \mathbf{0}\}. \quad (3.4.6)$$

由 (3.4.5) 和 (3.4.3) 易知集合 $G(\text{dom}F) + A \subset R^{q+1}$ 非空. 我们指出, $G(\text{dom}F) + A$ 还是凸的. 事实上, 任取 $(y^1, \mathbf{z}^1), (y^2, \mathbf{z}^2) \in G(\text{dom}F) + A$, 由 (3.4.5) 和 (3.4.6) 可知存在 $\mathbf{x}^1, \mathbf{x}^2 \in \text{dom}F$, 使

$$F(\mathbf{x}^1) \leq y^1, F(\mathbf{x}^2) \leq y^2,$$

$$\mathbf{h}(\mathbf{x}^1) = \mathbf{z}^1, \mathbf{h}(\mathbf{x}^2) = \mathbf{z}^2.$$

于是,对任意的 $\lambda \in (0, 1)$, 由于 f 是凸函数和 $h_j (j = 1, \dots, q)$ 是仿射函数, 我们有

$$\begin{aligned} F(\lambda x^1 + (1 - \lambda)x^2) &\leq \lambda F(x^1) + (1 - \lambda)F(x^2) \\ &\leq \lambda y^1 + (1 - \lambda)y^2, \\ h(\lambda x^1 + (1 - \lambda)x^2) &\leq \lambda h(x^1) + (1 - \lambda)h(x^2) \\ &\leq \lambda z^1 + (1 - \lambda)z^2. \end{aligned}$$

据此, 得到 $(\lambda y^1 + (1 - \lambda)y^2, \lambda z^1 + (1 - \lambda)z^2) \in G(\text{dom}F) + A$, 故由定义 1.1.1 知 $G(\text{dom}F) + A$ 是凸集.

另一方面, 对于非空凸集

$$D = \{(y, z) \in R \times R^q \mid y < 0, z = \theta\}, \quad (3.4.7)$$

由 (3.4.3) 和 (3.4.6) 可知

$$D \not\subset G(\text{dom}F) + A.$$

由于任何有限维非空凸集的相对内部是非空的, 故有

$$(G(\text{dom}F) + A)^{\circ} \neq \emptyset, D^{\circ} \neq \emptyset.$$

因此, 凸集 $G(\text{dom}F) + A$ 和 D 是可分离的. 根据定理 1.3.5 的 (1) 可知, 存在 $(w, \mu) \in R \times R^q \setminus \{\theta\}$, 使得

$$\begin{aligned} \langle w, y' \rangle + \langle \mu, z' \rangle &\leq \langle w, y'' \rangle + \langle \mu, z'' \rangle \\ \forall (y', z') \in D \quad \forall (y'', z'') \in G(\text{dom}F) + A. \end{aligned}$$

记 $\mu = (\mu_1, \dots, \mu_q)^T \in R^q$, 在上式中取 $y'' = F(x)$, $z'' = h(x)$ ($x \in \text{dom}F$), 注意 $y' < 0$ 和 $z' = \theta$ (见 (3.4.7)), 则有

$$\begin{aligned} \langle w, y' \rangle &\leq \langle w, F(x) \rangle + \langle \mu, h(x) \rangle \\ \forall y' \in (-\infty, 0), \forall x \in \text{dom}F, \end{aligned} \quad (3.4.8)$$

即

$$wy' \leq wF(x) + \sum_{j=1}^q \mu_j h_j(x)$$

$$\forall y' \in (-\infty, 0), \forall x \in \text{dom} F.$$

令 $y' \rightarrow 0$, 得到

$$wF(x) + \sum_{j=1}^q \mu_j h_j(x) \geq 0 \quad \forall x \in \text{dom} F,$$

这就是(3.4.4). 下面说明 $w > 0$. 首先, 因为(3.4.8)中的 y' 可趋于 $-\infty$ 而其右端是有限数, 故 w 不能取负值. 此外, 由于 $\theta \in (h(\text{dom} F))'$, 则知 $w \neq 0$. 因为否则, 若 $w = 0$, 则由(3.4.8), 将有

$$\langle \mu, h(x) \rangle \geq 0 \quad \forall x \in \text{dom} F,$$

即

$$\langle \mu, z \rangle \geq 0 \quad \forall z \in h(\text{dom} F).$$

因为 $\theta \in (h(\text{dom} F))'$, 故上式仅当 $\mu = \theta$ 时才有可能, 这导致与 $(w, \mu) \neq \theta$ 相矛盾, 从而引理得证. \square

下面阐述约束凸规划问题(MP)的最优解要满足的最优性条件.

定理 3.4.5 设 $f: \mathcal{X} \rightarrow R \cup \{+\infty\}$ 和 $g_i: \mathcal{X} \rightarrow R \cup \{+\infty\} (i = 1, \dots, p)$ 是正常凸函数, $h_j: \mathcal{X} \rightarrow R (j = 1, \dots, q)$ 是连续的仿射函数, $X \neq \emptyset$ 由(3.4.1)确定(其中 $\mathcal{V} = \mathcal{X}$). 若 f 和 $g_i (i = 1, \dots, p)$ 在点 $\bar{x} \in X$ 处有限并且是连续的, 则 $\bar{x} \in \bar{X}$ 当且仅当存在 $\bar{w} \geq 0, (\bar{\lambda}_1, \dots, \bar{\lambda}_p)^T \in R_+^p$ 和 $(\bar{\mu}_1, \dots, \bar{\mu}_q)^T \in R^q$ 使得

$$\begin{cases} \theta \in \bar{w} \partial f(\bar{x}) + \sum_{i=1}^p \bar{\lambda}_i \partial g_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \partial h_j(\bar{x}); \\ \bar{\lambda}_i g_i(\bar{x}) = 0, i = 1, \dots, p; \\ (\bar{w}, \bar{\lambda}_1, \dots, \bar{\lambda}_p)^T \neq \theta. \end{cases} \quad (3.4.9)$$

证明 记 $h = (h_1, \dots, h_q)^T$, 先设 $h: \mathcal{X} \rightarrow R^q$ 是满射的, 即 $h(\mathcal{X}) = R^q$ [19]. 因为 f 和 $g_i (i = 1, \dots, p)$ 在点 \bar{x} 处有限且是连续

的,故有

$$\bar{x} \in (\text{dom} f)^i \cap \left(\bigcap_{i=1}^p (\text{dom} g_i)^i \right).$$

从 $\bar{x} \in X$, 由 (3.4.1) 知 $h(\bar{x}) = \theta$, 因为 h 是仿射向量函数和满射的, 得知

$$\theta = h(\bar{x}) \in \left[h(\text{dom} f) \cap \left(\bigcap_{i=1}^p (\text{dom} g_i) \right) \right]^i. \quad (3.4.10)$$

由 $\bar{x} \in \bar{X}$, 根据定义 3.4.1 有

$$f(x) - f(\bar{x}) \geq 0 \quad \forall x \in \mathcal{X},$$

按 (3.4.1) (当 $\mathcal{V} = \mathcal{X}$) 得

$$\begin{aligned} \max_{1 \leq i \leq p} [f(x) - f(\bar{x}), g_i(x)] &\geq 0 \\ \forall x \in \{x \in \mathcal{X} \mid h(x) = \theta\}. \end{aligned} \quad (3.4.11)$$

令

$$F(x) = \max_{1 \leq i \leq p} [f(x) - f(\bar{x}), g_i(x)], \quad (3.4.12)$$

由 (3.4.11) 我们有

$$F(x) \geq 0 \quad \forall x \in \{x \in \mathcal{X} \mid h(x) = \theta\}, \quad (3.4.13)$$

并且

$$\text{dom} F = \text{dom} f \cap \left(\bigcap_{i=1}^p (\text{dom} g_i) \right).$$

此外, 从 (3.4.10) 知

$$\theta \in (h(\text{dom} F)),$$

于是由引理 3.4.4 得知 (3.4.13) 成立的充分必要条件是, 存在 $w > 0$ 和 $\tilde{\mu}_j (j = 1, \dots, q)$, 使得

$$wF(x) + \sum_{j=1}^q \tilde{\mu}_j h_j(x) \geq 0 \quad \forall x \in \text{dom} F.$$

由于 $w > 0$, 上式等价于

$$wF(x) + \sum_{j=1}^q \tilde{\mu}_j h_j(x) \geq 0 \quad \forall x \in \mathcal{X}. \quad (3.4.14)$$

因为 $E(x) = wF(x) + \sum_{j=1}^q \tilde{\mu}_j h_j(x)$ 也是 \mathcal{X} 上的正常凸函数, 由 (3.4.13) 可知 (3.4.14) 意味着它在 $\bar{x} \in \mathcal{X}$ 处达到最优值 $E(\bar{x}) = 0$, 所以根据定理 3.4.1, 有

$$0 \in \partial E(\bar{x}) = \partial \left[wF + \sum_{j=1}^q \tilde{\mu}_j h_j \right](\bar{x}).$$

据此, 由定理 3.3.4, 从上式得到

$$0 \in w\partial F(\bar{x}) + \sum_{j=1}^q \tilde{\mu}_j \partial h_j(\bar{x}). \quad (3.4.15)$$

另外, 由 (3.4.12) 根据定理 3.3.3 有

$$\partial F(\bar{x}) = \text{co} \left\{ \partial f(\bar{x}), \bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x}) \right\},$$

其中

$$I(\bar{x}) = \{i \in \{1, \dots, p\} \mid g_i(\bar{x}) = 0\}. \quad (3.4.16)$$

因此, 由定理 1.1.8 知, 存在

$$\lambda_0 \geq 0, \lambda_i \geq 0 \ (i \in I(\bar{x})), \lambda_i = 0 \ (i \notin I(\bar{x})), \sum_{i=0}^p \lambda_i = 1, \quad (3.4.17)$$

使得

$$\partial F(\bar{x}) = \lambda_0 \partial f(\bar{x}) + \sum_{i=1}^p \lambda_i \partial g_i(\bar{x}).$$

将它代入 (3.4.15) 有

$$0 \in w\lambda_0 \partial f(\bar{x}) + w \sum_{i=1}^p \lambda_i \partial g_i(\bar{x}) + \sum_{j=1}^q \tilde{\mu}_j \partial h_j(\bar{x}).$$

令 $\tilde{w} = w\lambda_i$, $\tilde{\lambda}_i = w\lambda_i (i = 1, \dots, p)$, 从上式即得(3.4.9)的第1式. 再由(3.4.16)和(3.4.17)可得(3.4.9)的第2式和第3式.

以上过程是可逆的, 故所得结果是 $\bar{x} \in \tilde{X}$ 的充分必要条件.

若 $h: \mathcal{X} \rightarrow R^q$ 不是满射的, 则可取 $h(\mathcal{X}) \neq \emptyset$ 来替代 R^q , 作同样的讨论, 即可使本定理得证. \square

定理 3.4.5 称为凸规划的 Fritz John 定理, 其中的(3.4.9)称为凸规划问题(MP)的最优解的 Fritz John 条件.

推论 3.4.6 设 $f: R^n \rightarrow R \cup \{+\infty\}$ 和 $g_i: R^n \rightarrow R \cup \{+\infty\} (i = 1, \dots, p)$ 是正常凸函数, $h_j: R^n \rightarrow R (j = 1, \dots, q)$ 是线性函数, $X \neq \emptyset$ 由(3.4.1)确定(其中 $\mathcal{X} = R^n$). 若 f 和 $g_i (i = 1, \dots, p)$ 在点 $\bar{x} \in X$ 处可微, 则 $\bar{x} \in \tilde{X}$ 当且仅当存在 $\tilde{w} \geq 0$, $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_p)^T \in R^p$ 和 $(\tilde{\mu}_1, \dots, \tilde{\mu}_q)^T \in R^q$, 使得

$$\begin{cases} \tilde{w} \nabla f(\bar{x}) + \sum_{i=1}^p \tilde{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^q \tilde{\mu}_j \nabla h_j(\bar{x}) = 0; \\ \tilde{\lambda}_i g_i(\bar{x}) = 0, i = 1, \dots, p; \\ (\tilde{w}, \tilde{\lambda}_1, \dots, \tilde{\lambda}_p)^T \neq \mathbf{0}. \end{cases}$$

证明 由推论 3.2.5, 这时 $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$, $\partial g_i(\bar{x}) = \{\nabla g_i(\bar{x})\} (i = 1, \dots, p)$ 和 $\partial h_j(\bar{x}) = \{\nabla h_j(\bar{x})\} (j = 1, \dots, q)$. 因此, 根据定理 3.4.5 即可得证. \square

最后, 介绍著名的 Kuhn-Tucker 条件.

定义 3.4.2 设 $g_i: \mathcal{X} \rightarrow R \cup \{\pm\infty\} (i = 1, \dots, p)$ 和 $h_j: \mathcal{X} \rightarrow R \cup \{\pm\infty\} (j = 1, \dots, q)$ 是广义实值函数, X 由(3.4.1)确定. 若存在点 $\bar{x} \in X$, 使得

$$g_i(\bar{x}) < 0, i = 1, \dots, p; h_j(\bar{x}) = 0, j = 1, \dots, q, \quad (3.4.18)$$

则称(MP)的约束集 X (或约束函数 $g_i (i = 1, \dots, p)$, $h_j (j = 1, \dots, q)$) 满足 Slater 约束规格.

定理 3.4.7 设 $f: \mathcal{K} \rightarrow R \cup \{+\infty\}$ 和 $g_i: \mathcal{K} \rightarrow R \cup \{+\infty\} (i=1, \dots, p)$ 是正常凸函数, $h_j: \mathcal{K} \rightarrow R (j=1, \dots, q)$ 是连续的仿射函数, $X \neq \emptyset$ 由 (3.4.1) 确定 (其中 $\mathcal{V} = \mathcal{K}$). 若 f 和 $g_i (i=1, \dots, p)$ 在点 $\bar{x} \in X$ 处有限并且是连续的, X 满足 Slater 约束规格, 则 $\bar{x} \in \tilde{X}$ 当且仅当存在 $(\bar{\lambda}_1, \dots, \bar{\lambda}_p)^T \in R_+^p$ 和 $(\bar{\mu}_1, \dots, \bar{\mu}_q)^T \in R^q$, 使得

$$\begin{cases} 0 \in \partial f(\bar{x}) + \sum_{i=1}^p \bar{\lambda}_i \partial g_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \partial h_j(\bar{x}); \\ \bar{\lambda}_i g_i(\bar{x}) = 0, i = 1, \dots, p. \end{cases} \quad (3.4.19)$$

证明 充分性. 由定理 3.4.5, 取 $\tilde{\omega} = 1$ 即得.

必要性. 根据定理 3.4.5, 只要在条件 (3.4.18) 下证明其中的 $\tilde{\omega} > 0$ 即可. 事实上, 若假设 $\tilde{\omega} = 0$, 则由 (3.4.9) 的第 1 式, 有

$$0 \in \sum_{i=1}^p \bar{\lambda}_i \partial g_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \partial h_j(\bar{x}),$$

由定理 3.3.4 即

$$0 \in \partial \left[\sum_{i=1}^p \bar{\lambda}_i g_i + \sum_{j=1}^q \bar{\mu}_j h_j \right](\bar{x}).$$

据此, 利用定理 3.4.1, 得知 $\bar{x} \in X$ 是以 $\sum_{i=1}^p \bar{\lambda}_i g_i(x) + \sum_{j=1}^q \bar{\mu}_j h_j(x)$ 为目标函数的无约束极小化问题 (P) 的最优解. 另外, 由 (3.4.18)、(3.4.16) 和 (3.4.17) 得知, 存在点 $\bar{x} \in X$ 使得

$$\sum_{i=1}^p \bar{\lambda}_i g_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j h_j(\bar{x}) < 0 = \sum_{i=1}^p \bar{\lambda}_i g_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j h_j(\bar{x}),$$

这导致与 \bar{x} 是 (P) 的最优解相矛盾. \square

定理 3.4.7 称为凸规划的 Kuhn-Tucker 定理, 其中的 (3.4.19) 称为凸规划问题 (MP) 的最优解的 Kuhn-Tucker 条件.

推论 3.4.8 设 $f: R^n \rightarrow R$ 和 $g_i: R^n \rightarrow R (i=1, \dots, p)$ 是

正常凸函数, $h_j: R^n \rightarrow R$ ($j = 1, \dots, q$) 是线性函数, $X \neq \emptyset$ 由 (3.4.1) 确定 (其中 $\mathcal{C} = R^n$), 并且 X 满足 Slater 约束规格, 则 $\bar{x} \in \bar{X}$ 当且仅当存在 $(\bar{\lambda}_1, \dots, \bar{\lambda}_p)^T \in R_+^p$ 和 $(\bar{\mu}_1, \dots, \bar{\mu}_q)^T \in R^q$, 使得

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^p \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \nabla h_j(\bar{x}) = 0; \\ \bar{\lambda}_i g_i(\bar{x}) = 0, i = 1, \dots, p. \end{cases}$$

证明 由定理 3.4.7 和推论 3.2.5 即可得证. \square

第 4 章 Lipschitz 函数的 G-次微分

Lipschitz 函数是一类重要的非光滑函数, 它的 G -次微分是目前各种意义下次微分中最具有代表性的. Lipschitz 函数的 G -次微分具有一般可微函数的许多有用的性质, 如次可加性、乘法规则、复合函数链式法则以及中值定理等. 此外, 在有限维空间中的 Lipschitz 函数是几乎处处可微的, 而对于凸的 Lipschitz 函数而言, 则它几乎就是可微的.

本章将先介绍 Lipschitz 函数的 G -方向导数和 G -次微分概念. 然后, 讨论它们的有关性质、几何特性以及在有限维空间的某些特殊结果.

§ 4.1 G -方向导数

设 \mathcal{B} 是 Banach 空间, $x \in \mathcal{B}$, $\|x\|$ (或 $\|x\|_{\mathcal{B}}$) 表示 x (在 \mathcal{B} 中) 的范数, B 表示 (\mathcal{B} 中的) 开单位球, 即

$$B = \{x \in \mathcal{B} \mid \|x\| < 1\}.$$

定义 4.1.1 设集合 $S \subset \mathcal{B}$ 非空, $f: S \rightarrow R$ 是实值函数.

(1) 若对任意的 $x^1, x^2 \in S$ 和某常数 $\gamma > 0$ 有

$$|f(x^1) - f(x^2)| \leqslant \gamma \|x^1 - x^2\|,$$

则称函数 f 在 S 上满足 (γ 秩) Lipschitz 条件, γ 称为是 f 的 Lipschitz 常数.

(2) 若对点 $x^0 \in S$, 存在 $\varepsilon > 0$, 使得 f 在 $x^0 + \varepsilon B$ 上满足 (γ 秩) Lipschitz 条件, 则称 f 在点 x^0 附近满足 (γ 秩) Lipschitz 条件, 或 f 在点 x^0 处是 (γ 秩) Lipschitz 的. 若对任意的 $x \in S$, f 在点 x 附近都满足 (γ 秩) Lipschitz 条件, 则称 f 是 S 上的 (γ 秩) Lipschitz 函数 (泛函), 或 f 在 S 上是 (γ 秩) Lipschitz 的.

Lipschitz 函数类包含了很大一类函数. 例如, 凸函数就是其有效域上的 Lipschitz 函数, 可微函数都是 Lipschitz 函数.

定义 4.1.2 设集合 $S \subset \mathcal{B}$ 非空, 函数 $f: S \rightarrow R$ 在点 $x^0 \in S$ 处是 (γ 秩) Lipschitz 的, $d \in \mathcal{B}$. 若极限

$$f^0(x^0; d) = \limsup_{\substack{x \rightarrow x^0 \\ t \rightarrow 0^+}} \frac{f(x + td) - f(x)}{t} \quad (4.1.1)$$

存在, 则称它是 f 在点 x^0 处沿方向 d 的广义方向导数或 G -方向导数.

定理 4.1.1 设集合 $S \subset \mathcal{B}$ 非空, f 在 $x \in S$ 处是 (γ 秩) Lipschitz 的.

(1) $f^0(x; d)$ 关于 d 是正齐次和次可加的 (即次线性的), 有限的, 并且满足

$$|f^0(x; d)| \leq \gamma \|d\|.$$

$$(2) f^0(x; -d) = -f^0(x; d).$$

证明 (1) 由 f 在 x 附近满足 γ 秩 Lipschitz 条件, 根据定义 4.1.2 可知, 对任意的 $\alpha \geq 0$ 有

$$f^0(x; \alpha d) = \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + \alpha td) - f(x')}{t} = \alpha f^0(x; d),$$

故正齐次性成立. 又对任何 $d^1, d^2 \in \mathcal{B}$ 有

$$f^0(x; d^1 + d^2) = \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td^1 + td^2) - f(x')}{t}$$

$$\begin{aligned} &\leq \lim_{\substack{x' + td^2 \rightarrow x \\ t \rightarrow 0^+}} \sup \frac{f(x' + td^1 + td^2) - f(x' + td^2)}{t} \\ &\quad + \lim_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \sup \frac{f(x' + td^2) - f(x')}{t}, \end{aligned}$$

即

$$f^0(x; d^1 + d^2) \leq f^0(x; d^1) + f^0(x; d^2),$$

故它具有次可加性. 由此, $f^0(x; d)$ 关于 d 是次线性泛函.

另外, 按定义 4.1.2 和定义 4.1.1 的(1), 有

$$|f^0(x; d)| \leq \frac{|f(x' + td) - f(x')|}{t} \leq \gamma \|d\|,$$

因此, 对于任给定的 d , $|f^0(x; d)|$ 是有限的.

(2) 由定义 4.1.2 知

$$f^0(x; -d) = \lim_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \sup \frac{f(x' - td) - f(x')}{t},$$

令 $x'' = x' - td$, 从上式有

$$f^0(x; -d) = - \lim_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \sup \frac{f(x'' + td) - f(x'')}{t},$$

故 $f^0(x; -d) = -f^0(x; d)$. \square

注 4.1.1 由定理 4.1.1 的(1)和(2), 易知有

$$f^0(x; ad) = af^0(x; d), \quad a \in R. \quad (4.1.2)$$

定理 4.1.2 设集合 $S \subset \mathcal{B}$ 非空, f 在 $x \in S$ 处是 (γ 秩) Lipschitz 的.

(1) 函数 $f^0(x; d)$ 在点 (x, d) 处是上半连续的.

(2) $f^0(x; d)$ 关于 d 在 \mathcal{B} 上是 γ 秩 Lipschitz 的.

证明 (1) 考虑任意点列 $\{x_k\}$, $x_k \rightarrow x$, 以及 $\{d^k\} \subset \mathcal{B}$, $d^k \rightarrow d$ ($k \rightarrow \infty$). 对于每个 k , 由定义 4.1.2 知存在 $x'_k \in \mathcal{B}$ 和 $t_k > 0$

使得

$$\begin{aligned} \|x'_k - x_k\| + t_k &\leq \frac{1}{k}, \\ f^0(x_k; d^k) - \frac{1}{k} &\leq \frac{f(x'_k + t_k d^k) - f(x'_k)}{t_k} \\ &= \frac{f(x'_k + t_k d) - f(x'_k)}{t_k} \\ &\quad + \frac{f(x'_k + t_k d^k) - f(x'_k + t_k d)}{t_k}, \end{aligned}$$

并且上式右端第 2 式小于 $\gamma \|d^k - d\|$. 令 $k \rightarrow \infty$, 从上式即有

$$\limsup_{k \rightarrow \infty} f^0(x_k; d^k) \leq f^0(x; d).$$

由此, 根据注 2.1.3 即知 $f^0(x; d)$ 在 (x, d) 处是上半连续的.

(2) 对于任意给定的 $d^1, d^2 \in \mathcal{B}$, 我们有

$$f(x' + td^1) - f(x') \leq f(x' + td^2) - f(x') + \gamma \|d^1 - d^2\| t,$$

从而

$$\begin{aligned} \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td^1) - f(x')}{t} &\leq \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td^2) - f(x')}{t} \\ &\quad + \gamma \|d^1 - d^2\|. \end{aligned}$$

据此, 由定义 4.1.2 得

$$f^0(x; d^1) \leq f^0(x; d^2) + \gamma \|d^1 - d^2\|.$$

由 $d^1, d^2 \in \mathcal{B}$ 的任意性, 按定义 4.1.1 知 $f^0(x; d)$ 关于 d 在 \mathcal{B} 上是 γ 秩 Lipschitz 的. \square

§ 4.2 G-次微分

设 \mathcal{B} 是 Banach 空间, \mathcal{B}^* 是 \mathcal{B} 的对偶空间, $x^* \in \mathcal{B}^*$ 是 $x \in$

\mathcal{B} 的连续线性泛函, $\langle x^*, x \rangle$ 表示泛函 x^* 在 x 处的值.

定义 4.2.1 设集合 $S \subset \mathcal{B}$ 非空, 函数 $f: S \rightarrow R$ 在点 $x^0 \in S$ 处是 (γ 秩) Lipschitz 的. 若 $x^* \in \mathcal{B}^*$, 并且

$$\langle x^*, d \rangle \leq f^0(x^0; d) \quad \forall d \in \mathcal{B}, \quad (4.2.1)$$

则称 x^* 是 f 在点 x^0 处的广义次梯度或 G -次梯度. f 在点 x^0 处的所有 G -次梯度的集合称为 f 在点 x^0 处的广义次微分或 G -次微分, 记作 $\partial^0 f(x^0)$, 即

$$\partial^0 f(x^0) = \{x^* \in \mathcal{B}^* \mid \langle x^*, d \rangle \leq f^0(x^0; d) \quad \forall d \in \mathcal{B}\}. \quad (4.2.2)$$

若 $\partial^0 f(x^0) \neq \emptyset$, 则称 f 在点 x^0 处是广义次可微的或 G -次可微的. 若 f 在 S 的每一点处是 G -次可微的, 则称 f 在 S 上是 G -次可微的.

例 4.2.1 设 $\mathcal{B} = R$, $f(x) = |x|$. 按定义 4.1.2, 对任意的 $d \in R$ 有 $f^0(x; d) = d$. 由定义 4.2.1, 得

$$\partial^0 f(x) = \begin{cases} 1, & x > 0; \\ [-1, 1], & x = 0; \\ -1, & x < 0. \end{cases}$$

定理 4.2.1 设集合 $S \subset \mathcal{B}$ 非空, f 在 $x \in S$ 处是 γ 秩 Lipschitz 的.

- (1) $\partial^0 f(x)$ 是非空凸集, 且在 \mathcal{B}^* 中是紧的 (在弱 * 意义下).
- (2) 对任意的 $d \in \mathcal{B}$, 有

$$f^0(x; d) = \max \{\langle x^*, d \rangle \mid x^* \in \partial^0 f(x)\}.$$

证明 (1) 由定理 4.1.1 和 Hahn-Banach 定理^[21], 至少存在一线性泛函 $x^*: \mathcal{B} \rightarrow R$, 使得对任意的 $d \in \mathcal{B}$ 有 $\langle x^*, d \rangle \leq f^0(x; d)$. 因此, $\partial^0 f(x)$ 是非空的. $\partial^0 f(x)$ 显然是凸的. 此外, 利用定理 4.1.1 的 (1), 有

$$\langle x^*, d \rangle \leq f^0(x; d) \leq \gamma \|d\|.$$

由 Alaoglu 定理^[20]即知在弱*意义下 $\partial^0 f(x)$ 是紧的.

(2) 反证: 假设对某个 $d^0 \in \mathcal{B}$, 有

$$f^0(x; d^0) > \langle x^*, d^0 \rangle \quad \forall x^* \in \partial^0 f(x).$$

由 Hahn-Banach 定理可知, 存在连续线性泛函 $\bar{x}^* \in \mathcal{B}^*$, 使得

$$f^0(x; d) \geq \langle \bar{x}^*, d \rangle \quad \forall d \in \mathcal{B},$$

并且 $f^0(x; d^0) = \langle \bar{x}^*, d^0 \rangle$. 然而 $\bar{x}^* \in \partial^0 f(x)$, 这显然是矛盾的. \square

定理 4.2.2 设集合 $S \subset \mathcal{B}$ 非空, f 在 $x \in S$ 处是 Lipschitz 的.

(1) $x^* \in \partial^0 f(x)$ 当且仅当

$$\langle x^*, d \rangle \leq f^0(x; d) \quad \forall d \in \mathcal{B}.$$

(2) 设 $\{x_k\} \subset \mathcal{B}$, $\{x_k^*\} \subset \mathcal{B}^*$, $x_k^* \in \partial^0 f(x_k)$ ($k = 1, 2, \dots$). 若 $x_k \rightarrow x$, $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$), 则 $x^* \in \partial^0 f(x)$.

(3) 设 $B \subset \mathcal{B}$ 是开单位球, 则

$$\partial^0 f(x) = \bigcap_{\delta > 0} \bigcup_{x' \in x + \delta B} \partial^0 f(x').$$

(4) 若 $B \subset R^n$, 则对任给的 $\varepsilon > 0$, 存在 $\delta > 0$ 使得

$$\partial f^0(x') \subset \partial f^0(x) + \varepsilon B \quad \forall x' \in x + \delta B.$$

证明 (1) 由定义 4.2.1 可得.

(2) 利用定理 4.1.2 的(1)和 $\langle x_k^*, d \rangle \leq f^0(x_k; d)$, 令 $k \rightarrow \infty$, 有

$$\langle x^*, d \rangle \leq f^0(x; d) \quad \forall d \in \mathcal{B}.$$

因此, 得知 $x^* \in \partial^0 f(x)$.

(3) 和 (4) 均可由 (2) 推得. \square

定理 4.2.3 设 $f: \mathcal{B} \rightarrow R$ 在 $\bar{x} \in \mathcal{B}$ 处是 Lipschitz 的. 若 \bar{x} 是

无约束极小化问题(UMP)的最优解, 则 $\mathbf{0} \in \partial^0 f(\bar{x})$.

证明 由定义 3.4.1 和定义 4.1.2 可知

$$f^0(\bar{x}; d) \geqslant 0 \quad \forall d \in \mathcal{B}.$$

由此, 按定义 4.2.1 得 $\mathbf{0} \in \partial^0 f(\bar{x})$. \square

下面讨论 G -次微分与其他意义下的导数的关系.

设 \mathcal{D} 是 Banach 空间.

定义 4.2.2 设 $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是映射, 点 $x^0 \in \mathcal{B}$. 若对 x^0 附近的任意点 x^1, x^2 和某常数 $\gamma > 0$ 有

$$\|\varphi(x^1) - \varphi(x^2)\|_{\mathcal{D}} \leqslant \gamma \|x^1 - x^2\|_{\mathcal{B}},$$

则称映射 φ 在点 x^0 附近满足 (γ 秩) Lipschitz 条件, 或在点 x^0 处是 Lipschitz 的.

设 $L(\mathcal{B}, \mathcal{D})$ 表示 \mathcal{B} 到 \mathcal{D} 的所有连续线性映射组成的集合.

定义 4.2.3 设 $S \subset \mathcal{B}$ 是紧集, $\varphi: S \rightarrow \mathcal{D}$ 是映射, 点 $x^0 \in \mathcal{B}$, $D_S \varphi(x^0) \in L(\mathcal{B}, \mathcal{D})$. 若对任意的 $d \in \mathcal{B}$, 有

$$\lim_{\substack{x \rightarrow x^0 \\ t \rightarrow 0^+}} \frac{\varphi(x+td) - \varphi(x)}{t} = \langle D_S \varphi(x^0), d \rangle \quad (4.2.3)$$

成立, 并且左端的极限关于 d 一致收敛, 则称 φ 在点 x^0 处是严格可微的, 并且称 $D_S \varphi(x^0)$ 是 φ 在点 x^0 处的严格导数. 若 φ 在 S 上的每一点处是严格可微的, 则称 φ 是 S 上的严格可微映射, 或在 S 上是严格可微的.

定理 4.2.4 设 $S \subset \mathcal{B}$ 是紧集, $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是映射, $x^* \in L(\mathcal{B}, \mathcal{D})$. φ 在 $x \in S$ 处是严格可微的且 $D_S \varphi(x) = x^*$ 的充要条件是: φ 在 x 附近满足 Lipschitz 条件, 并且对任意的 $d \in \mathcal{B}$ 有

$$\lim_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{\varphi(x' + td) - \varphi(x')}{t} = \langle x^*, d \rangle. \quad (4.2.4)$$

证明 必要性. 由 φ 在 x 处是严格可微的及 $D_S \varphi(x) = x^*$, 按定义 4.2.3 可知, 对任意的 $d \in \mathcal{B}$, (4.2.4) 成立. 下面证明 φ 在 x

附近满足 Lipschitz 条件. 反证: 假设不然, 则存在点列 $\{x_k\}$, $x_k \rightarrow x$, $\{x'_k\}$, $x'_k \rightarrow x'$ ($k \rightarrow \infty$), 使得对每个 k 有 $x_k, x'_k \in x + (1/k)B$ 和

$$\|\varphi(x'_k) - \varphi(x_k)\|_{\mathcal{B}} \geq k \|x'_k - x_k\|_{\mathcal{B}}.$$

再定义序列 $\{t_k\}$, $t_k > 0$, 以及 $\{d^k\}$ 使得 $x'_k = x_k + t_k d^k$ 和 $\|d^k\| = k^{-\frac{1}{2}}$, $t_k \rightarrow 0$ ($k \rightarrow \infty$). 取集合 $A = \{d^k\} \cup \{\theta\}$, 显然它是紧集. 因此, 根据定义 4.2.3, 对任给的 $\varepsilon > 0$, 存在 N 使得对任何 $k \geq N$ 和所有的 $d \in A$ 有

$$\left\| \frac{\varphi(x'_k + t_k d) - \varphi(x'_k)}{t_k} - \langle D_S \varphi(x), d \rangle \right\|_{\mathcal{B}} \leq \varepsilon. \quad (4.2.5)$$

但是, 另一方面有

$$\left\| \frac{\varphi(x'_k + t_k d) - \varphi(x'_k)}{t_k} - \langle D_S \varphi(x), d \rangle \right\|_{\mathcal{B}} > k^{\frac{1}{2}},$$

令 $k \rightarrow \infty$, 这导致与 (4.2.5) 相矛盾.

充分性. 我们证明 φ 在 x 处是严格可微的. 设 $S \subset \mathcal{B}$ 是任一紧子集, $\varepsilon > 0$ 是任给的数. 对任意的 $d \in \mathcal{B}$ 存在正数 $\delta(d)$, 使得对任意的 $x' \in x + \varepsilon B$ 和 $t \in (0, \delta)$ 有

$$\left\| \frac{\varphi(x' + td) - \varphi(x')}{t} - \langle D_S \varphi(x), d \rangle \right\|_{\mathcal{B}} < \varepsilon. \quad (4.2.6)$$

因为 φ 在 x 处是 Lipschitz 的, 故存在 δ' 使得对任意的 $d' \in d + \delta' B$ 和 $t \in (0, \delta')$ 有 $x' + td' \in x + \delta' B$, 以及

$$\|\varphi(x' + td') - \varphi(x' + td)\|_{\mathcal{B}} < \gamma t \|d' - d\|_{\mathcal{B}}. \quad (4.2.7)$$

由 (4.2.6) 和 (4.2.7) 可重新选择 $\delta(d)$, 使得对任意的 $d' \in d + \delta B$ 及任意的 $x' \in x + \delta B$ 和 $t \in (0, \delta)$, 有

$$\left\| \frac{\varphi(x' + td') - \varphi(x')}{t} - \langle D_S \varphi(x), d \rangle \right\|_{\mathcal{B}} < 2\varepsilon.$$

由此,得到 $\{d + \delta(d)B | d \in \mathcal{B}\}$ 是 S 的一个开覆盖. 由于 S 是紧的, 则它存在一个有限子覆盖 $d_1 + \delta(d_1)B, \dots, d_n + \delta(d_n)B$. 选择 $\delta = \min_{1 \leq k \leq n} \delta(d_k)$, 则对任意的 $d \in S$ 及任意的 $x' \in x + \delta'B$ 和 $t \in (0, \delta')$, (4.2.6) 成立. 因此, 由定义 4.2.3, φ 在 x 处是严格可微的. \square

以下考虑 G -次微分与 Frechet 导数、Gâteaux 导数以及严格导数之间的关系.

定理 4.2.5 设集合 $S \subset \mathcal{B}$ 非空, $f: S \rightarrow R$ 在 $x \in S$ 处是 Lipschitz 的, 并且 f 在 x 处是 Frechet 可微的.

(1) 设 $f'(x)$ 是 f 在 x 处的 Frechet 导数, 则 $f'(x) \in \partial^0 f(x)$.

(2) 设 f'_x 是 f 在 x 处的 Gâteaux 导数, 则 $f'_x \in \partial^0 f(x)$.

证明 (1) 因为 f 在 x 处是 Frechet 可微的, 由定理 3.1.5 和注 3.1.1 得知

$$\langle f'(x), d \rangle = f'(x; d) \quad \forall d \in \mathcal{B}.$$

又由定义 3.1.1 的(3)和定义 4.1.2, 有

$$f'(x; d) \leq f^0(x; d) \quad \forall d \in \mathcal{B},$$

所以

$$\langle f'(x), d \rangle \leq f^0(x; d) \quad \forall d \in \mathcal{B}.$$

据此, 按定义 4.2.1 即知 $f'(x) \in \partial^0 f(x)$.

(2) 由定理 3.1.5 知 $f'(x) = f'_x$, 根据(1)即得证. \square

定理 4.2.6 设集合 $S \subset \mathcal{B}$ 非空, 则 $f: S \rightarrow R$ 在 $x \in S$ 处是严格可微的当且仅当 f 在 x 处是 Lipschitz 的, 并且 $\partial^0 f(x)$ 是单点集, 即 $\partial^0 f(x) = \{D_S f(x)\}$.

证明 必要性. 由于 f 在 x 处是严格可微的, 从定理 4.2.4 知 f 在 x 处是 Lipschitz 的. 据此, 按定义 4.1.2 和定义 4.2.3 知, 对任意的 $d \in \mathcal{B}$ 有 $f^0(x; d) = \langle D_S f(x), d \rangle$. 再由定理 4.2.1 的(2)即得 $\partial^0 f(x) = \{D_S f(x)\}$.

充分性. 根据定理 4.2.1 的(2), 只要证明对任意的 $d \in \mathcal{B}$ 有

(4.2.4) 成立 ($\varphi = f$), 即可推出 f 在 x 处是严格可微的. 现在证明对任意的 $d \in \mathscr{B}$ 有

$$f^0(x; d) = \langle x^*, d \rangle, \quad (4.2.8)$$

其中 $\partial^0 f(x) = \{x^*\}$. 反证: 假设 (4.2.8) 不成立, 则存在 $\bar{d} \in \mathscr{B}$ 使得 $f^0(x; \bar{d}) > \langle x^*, \bar{d} \rangle$. 由 Hahn-Banach 定理可知, 存在一点 $\bar{x}^* \in \mathscr{B}^*$, 使得 $f^0(x; d) \geq \langle \bar{x}^*, d \rangle$ 对任何 $d \in \mathscr{B}$ 成立. 特别是, 有 $f^0(x; \bar{d}) = \langle \bar{x}^*, \bar{d} \rangle$, 故 $\bar{x}^* \in \partial^0 f(x)$. 因为已知 $\partial^0 f(x)$ 是单点集, 所以 $\bar{x}^* = x^*$, 从而导致 $f^0(x; \bar{d}) > f^0(x; \bar{d})$, 产生矛盾. 从 (4.2.8), 我们有

$$\begin{aligned} & \liminf_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td) - f(x')}{t} \\ &= - \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x') - f(x' + td)}{t} \\ &= - \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td - td) - f(x' + td)}{t} \\ &= - f^0(x; -d) = - \langle \bar{x}^*, -d \rangle \\ &= \langle \bar{x}^*, d \rangle = f^0(x; d) \\ &= - \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f(x' + td) - f(x')}{t}. \end{aligned}$$

于是, 对于任意的 $d \in \mathscr{B}$, (4.2.4) 成立 ($\varphi = f$). \square

推论 4.2.7 设 $f: R^n \rightarrow R$ 在 $x \in R^n$ 处是 Lipschitz 的, $\varepsilon > 0$, 则对任意的 $x' \in x + \varepsilon B$, $\partial^0 f(x')$ 是单点集当且仅当 f 在 $x + \varepsilon B$ 上是连续可微的.

证明 由定理 4.2.6 可得证. \square

关于 Lipschitz 函数的 G -次微分与凸函数的次微分之间有以下关系.

定理 4.2.8 设 $S \subset \mathscr{B}$ 是开凸集, $f: S \rightarrow R$ 是凸函数. 若 f 在 S 中一点的某邻域有界, 则对任意的 $x \in S$, f 在 x 处是 Lipschitz 的.

证明 首先证明对任意的 $x \in S$, f 在 x 的一个邻域有界. 不妨设 f 在点 θ 的邻域 $\epsilon B \subset S$ 上是有界的, 即有 $M \in R$ 使

$$f(x') \leq M \quad \forall x' \in \epsilon B. \quad (4.2.9)$$

由于 S 是开凸集, 可以选择 $\rho > 1$ 使 $y = \rho x \in S$. 令 $\lambda = 1/\rho$, 作集合

$$D = \{d \in \mathscr{B} \mid d = (1 - \lambda)x' + \lambda y, x' \in \epsilon B\},$$

则 $D = x + (1 - \lambda)\epsilon B$. 对任意的 $d \in D$, 由 (4.2.9) 和 f 是 S 上的凸函数, 有

$$f(d) \leq (1 - \lambda)f(x') + \lambda f(y) \leq M + \lambda f(y). \quad (4.2.10)$$

取 $d' \in D$ 使得 $x = \frac{d + d'}{2}$, 则有

$$f(x) \leq \frac{1}{2}f(d) + \frac{1}{2}f(d').$$

由 (4.2.10) 得

$$f(d) \geq 2f(x) - f(d') \geq 2f(x) - M - \lambda f(y). \quad (4.2.11)$$

从 (4.2.10) 和 (4.2.11) 知 f 在 x 的邻域有界, 取

$$N = \begin{cases} M + \lambda f(y), & f(x) \geq 0; \\ M + \lambda f(y) - 2f(x), & f(x) < 0, \end{cases}$$

故存在 $\delta > 0$, 使得

$$|f(v)| \leq N \quad \forall v \in x + 2\delta B. \quad (4.2.12)$$

对任意的 $x^1, x^2 \in x + 2\delta B$, 设 $x^3 = x^2 + \frac{\delta}{\alpha}(x^2 - x^1)$, 其中 $\alpha = \|x^2 - x^1\|$, 显然 $x^3 \in x + 2\delta B$. 因此, 我们有

$$x^2 = \frac{\delta}{\delta + \alpha}x^1 + \frac{\alpha}{\delta + \alpha}x^3.$$

因为 f 在 D 上是凸的, 所以

$$f(x^2) \leq \frac{\delta}{\delta + \alpha}f(x^1) + \frac{\alpha}{\delta + \alpha}f(x^3),$$

从而

$$\begin{aligned} f(x^2) - f(x^1) &\leq \frac{\delta}{\delta + \alpha}[f(x^3) - f(x^1)] \\ &\leq \frac{\alpha}{\delta}[f(x^3) - f(x^1)]. \end{aligned}$$

由 (4.2.12) 则得到

$$f(x^2) - f(x^1) \leq \frac{2N}{\delta} \|x^2 - x^1\|.$$

注意 x^1 和 x^2 是 $x + 2\delta B$ 中的任意点, 因此得知 f 在 x 处是 Lipschitz 的. \square

定理 4.2.9 设 $S \subset \mathcal{B}$ 是开凸集, 若 $f: S \rightarrow R$ 是凸函数, 并且 f 在 $x \in S$ 处是 Lipschitz 的, 则

$$f^0(x; d) = f'(x; d) \quad \forall d \in \mathcal{B}.$$

证明 对任给定的 $\delta > 0$, 我们有

$$f^0(x; d) = \lim_{\varepsilon \rightarrow 0^+} \sup_{\|x' - x\| \leq \varepsilon \delta} \sup_{0 < t < \varepsilon} \frac{f(x' + td) - f(x')}{t},$$

由 f 是凸函数可知, 映射: $t \rightarrow \frac{f(x' + td) - f(x')}{t}$ 关于 t 是非减函数. 因此, 有

$$f^0(x; d) = \lim_{\varepsilon \rightarrow 0^+} \sup_{\|x' - x\| \leq \varepsilon \delta} \frac{f(x' + \varepsilon d) - f(x')}{\varepsilon}.$$

再根据 f 在 $x' \in x + \varepsilon \delta B$ 上是 Lipschitz 的, 可知有 $\gamma > 0$ 使得

$$\left| \frac{f(x' + \varepsilon d) - f(x')}{\varepsilon} - \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} \right| \leq 2\gamma\delta,$$

故

$$\begin{aligned} f^0(x; d) &\leq \lim_{\varepsilon \rightarrow 0^+} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} + 2\delta\gamma \\ &= f'(x; d) + 2\delta\gamma. \end{aligned}$$

由 δ 的任意性推知 $f^0(x; d) \leq f'(x; d)$, 又显然有 $f'(x; d) \leq f^0(x; d)$, 因而 $f^0(x; d) = f'(x; d)$. \square

定理 4.2.9 表明, 对于在某点附近满足 Lipschitz 条件的凸函数, 它的次微分与 G-次微分是一致的.

定理 4.2.10 设集合 $S \subset \mathscr{B}$ 非空, $f: S \rightarrow R$ 在 $x \in S$ 处是 Lipschitz 的, 则 $f^0(x; d)$ 是 $\partial^0 f(x)$ 的支撑函数.

证明 由定义 4.2.1 和定义 1.3.4 即得证. \square

设集合 $T \subset \mathscr{B}^*$ 非空, 记 $\sigma_T^*: \mathscr{B} \rightarrow R \cup \{+\infty\}$ 是 T 的支撑函数.

定理 4.2.11 设 $S_1, S_2 \subset \mathscr{B}$ 是非空闭凸集, $T_1, T_2 \subset \mathscr{B}^*$ 是非空弱闭凸集.

(1) $S_1 \subset S_2$ 当且仅当 $\sigma_{S_1}(x^*) \leq \sigma_{S_2}(x^*)$ ($\forall x^* \in \mathscr{B}^*$).

(2) $T_1 \subset T_2$ 当且仅当 $\sigma_{T_1}^*(x) \leq \sigma_{T_2}^*(x)$ ($\forall x \in \mathscr{B}$).

(3) T_1 是弱*紧的当且仅当 $\sigma_{T_1}^*(x)$ 在 \mathscr{B} 上取有限值.

(4) 函数 $\sigma: \mathscr{B} \rightarrow R \cup \{+\infty\}$ 是正齐次和次可加的, 下半连续的, 以及不恒为 $+\infty$ 的充要条件是: 存在非空弱闭凸子集 $T \subset \mathscr{B}^*$ 使得 $\sigma = \sigma_T^*$, 并且 T 是唯一的.

证明 (1)和(2)由定义 1.3.2 可得证.

(3) 由 T_1 是弱*紧的, 对任意的 $x \in \mathscr{B}$, 从 $x^* \in T_1$ 知 $\langle x^*, x \rangle$ 为有限值. 按定义 1.3.4 得 $\sigma_{T_1}^*(x)$ 在 \mathscr{B} 上取有限值. 反之, 显然

成立.

(4) 充分性显然成立, 以下证明必要性. 设集合

$$T = \{x^* \in \mathcal{B}^* \mid \sigma(x) \geq \langle x^*, x \rangle, x \in \mathcal{B}\}.$$

首先证明 T 是凸集. 设对任意的 $x_1^*, x_2^* \in T$, 任意的 $\lambda \in (0, 1)$, 有

$$\sigma(x) \geq \langle x_1^*, x \rangle \quad \forall x \in \mathcal{B},$$

$$\sigma(x) \geq \langle x_2^*, x \rangle \quad \forall x \in \mathcal{B}.$$

因为 σ 是正齐次和次可加的, 则对任意的 $x \in \mathcal{B}$, 有

$$\begin{aligned} \sigma(x) &= \sigma(\lambda x + (1 - \lambda)x) \\ &\geq \lambda \sigma(x) + (1 - \lambda) \sigma(x) \\ &\geq \langle \lambda x_1^* + (1 - \lambda)x_2^*, x \rangle. \end{aligned}$$

由 T 的定义得 $\lambda x_1^* + (1 - \lambda)x_2^* \in T$, 故 T 是凸集. 为证明 T 是弱闭的, 设点列 $\{x_k^*\} \subset T$ ($k = 1, 2, \dots$), 对任意的 $x \in \mathcal{B}$ 有 $\langle x_k^*, x \rangle \rightarrow \langle x^*, x \rangle$ (弱收敛) ($k \rightarrow \infty$). 从

$$\sigma(x) \geq \langle x_k^*, x \rangle \quad \forall x \in \mathcal{B}, k = 1, 2, \dots,$$

得

$$\sigma(x) \geq \langle x^*, x \rangle \quad \forall x \in \mathcal{B},$$

则有 $x^* \in T$. T 的非空性的证明类似于定理 4.2.1 的(1)的证明. 下面证明

$$\sigma(x) = \sup_{x^* \in T} \langle x^*, x \rangle. \quad (4.2.13)$$

反证: 假设对某个 $x^0 \in \mathcal{B}$, 有

$$\sigma(x^0) > \langle x^*, x^0 \rangle \quad \forall x^* \in T.$$

根据 Hahn-Banach 定理, 存在一连续线性泛函 $\bar{x}^* \in \mathcal{B}^*$, 使得

$$\sigma(x) \geq \langle \bar{x}^*, x \rangle \quad \forall x \in \mathcal{B},$$

并且 $\sigma(\mathbf{x}^0) = \langle \bar{\mathbf{x}}^*, \mathbf{x}^0 \rangle$, 然而 $\bar{\mathbf{x}}^* \in T$, 这显然是矛盾的. 从而 (4.2.13) 得证. \square

最后, 介绍函数在一点处正则的概念.

定义 4.2.4 设 $f: \mathcal{B} \rightarrow R$ 是实值函数, 点 $\mathbf{x}^0 \in \mathcal{B}$. 若对任意的 $\mathbf{d} \in \mathcal{B}$, f 在点 \mathbf{x}^0 处沿 \mathbf{d} 的右方向导数 $f'_-(\mathbf{x}^0; \mathbf{d})$ 存在, 并且对任意的 $\mathbf{d} \in \mathcal{B}$, 有 $f^0(\mathbf{x}^0; \mathbf{d}) = f'_+(\mathbf{x}^0; \mathbf{d})$, 则称函数 f 在点 \mathbf{x}^0 处是正则的.

定理 4.2.12 设 $f: \mathcal{B} \rightarrow R$ 在 $\mathbf{x} \in \mathcal{B}$ 处是 Lipschitz 的.

(1) 若 f 在 \mathbf{x} 处是严格可微的, 则 f 在 \mathbf{x} 处是正则的.

(2) 若 f 是凸函数, 则 f 在 \mathbf{x} 处是正则的.

(3) 若 f 在 \mathbf{x} 处存在 Gâteaux 导数 f'_x , 并且 f 在 \mathbf{x} 处是正则的, 则 $\partial^0 f(\mathbf{x}) = \{f'_x\}$.

证明 (1) 由定义 4.2.3、定理 4.2.2 和定理 4.2.6, 有

$$\begin{aligned} f^0(\mathbf{x}; \mathbf{d}) &= \max \{ \langle \mathbf{x}^*, \mathbf{d} \rangle \mid \mathbf{x}^* \in \partial^0 f(\mathbf{x}) \} \\ &= \langle D_\delta f(\mathbf{x}), \mathbf{d} \rangle = f'(\mathbf{x}; \mathbf{d}). \end{aligned}$$

因此, 按定义 4.2.4 即得 f 在 \mathbf{x} 处是正则的.

(2) 由定理 4.2.9 可知, 对任意的 $\mathbf{d} \in \mathcal{B}$ 有 $f'(\mathbf{x}; \mathbf{d}) = f^0(\mathbf{x}; \mathbf{d})$.

(3) 由定义 4.2.4 可以推得. \square

定理 4.2.13 设 $f_1, f_2: \mathcal{B} \rightarrow R$ 在 $\mathbf{x} \in \mathcal{B}$ 处是正则的, $\alpha > 0, \beta > 0$, 则 $\alpha f_1 + \beta f_2$ 在 \mathbf{x} 处是正则的.

证明 因为 $f_1, f_2: \mathcal{B} \rightarrow R$ 在 \mathbf{x} 处是正则的, $\alpha > 0, \beta > 0$, 故对任意的 $\mathbf{d} \in \mathcal{B}$ 有

$$\begin{aligned} (\alpha f_1 + \beta f_2)^0(\mathbf{x}; \mathbf{d}) &\geq (\alpha f_1 + \beta f_2)'(\mathbf{x}; \mathbf{d}) \\ &= \alpha f'_1(\mathbf{x}; \mathbf{d}) + \beta f'_2(\mathbf{x}; \mathbf{d}) \\ &= \alpha f_1^0(\mathbf{x}; \mathbf{d}) + \beta f_2^0(\mathbf{x}; \mathbf{d}) \\ &\geq (\alpha f_1 + \beta f_2)^0(\mathbf{x}; \mathbf{d}). \end{aligned}$$

据此,有

$$(\alpha f_1 + \beta f_2)^0(x; d) = (\alpha f_1 + \beta f_2)'(x; d) \quad \forall d \in \mathcal{B},$$

于是由定义 4.2.4 即得证. \square

例 4.2.2 考虑函数 $f: R^n \rightarrow R, x = (x_1, \dots, x_m)^T$,

$$f(x) = f(x_1, \dots, x_m) = \max\{x_i | i = 1, \dots, m\}.$$

记 $I = \{i \in \{1, \dots, m\} | x_i = \max_{1 \leq j \leq m} x_j\}$, $d = (d_1, \dots, d_m)^T$, 则有 $f'(x; d) = \max_{i \in I} d_i$ 和

$$\partial^0 f(0) = \{(a_1, \dots, a_m)^T \in R^m | a_i \geq 0, i = 1, \dots, m;$$

$$a_i = 0, i \in I, \sum_{i=1}^m a_i = 1\}.$$

§ 4.3 G -次微分的性质

本节阐述 G -次微分的运算规则、中值定理以及复合函数的链式法则等性质. 我们指出, 以后引进的其他意义下的次微分, 都没有 G -次微分所具有的那样多好的性质.

设 \mathcal{B} 是 Banach 空间. 除非有特别的申明, 本节均设 $f: \mathcal{B} \rightarrow R$ 在点 $x \in \mathcal{B}$ 附近满足 Lipschitz 条件.

定理 4.3.1 设 $x \in \mathcal{B}, a \in R$, 则 $\partial^0(af)(x) = a\partial^0 f(x)$.

证明 当 $a \geq 0$ 时, 对任意的 $d \in \mathcal{B}$, 显然有

$$(af)^0(x; d) = af^0(x; d),$$

故 $\partial^0(af)(x) = a\partial^0 f(x)$. 现设 $x^* \in \partial^0(-f)(x)$, 则对任意的 $d \in \mathcal{B}$, 有 $(-f)^0(x; d) \geq \langle x^*, d \rangle$. 由定理 4.1.1 的(2), 对任意的 $d \in \mathcal{B}$ 有 $f^0(x; d) \geq \langle -x^*, d \rangle$, 因此得 $x^* \in -\partial^0 f(x)$. 反之, 从 $x^* \in -\partial^0 f(x)$ 不难得知有 $x^* \in \partial^0(-f)(x)$. \square

定理 4.3.2 设 $f_1, f_2: \mathcal{B} \rightarrow R$ 是实值函数. 若 f_1 和 f_2 在

$x \in \mathcal{B}$ 处是 Lipschitz 的, 则

$$\partial^0(f_1 + f_2)(x) \subset \partial^0 f_1(x) + \partial^0 f_2(x).$$

证明 易知 $f_1 + f_2$ 在 x 处是 Lipschitz 的, 由定义 4.1.2 得

$$\begin{aligned} & (f_1 + f_2)^0(x; d) \\ &= \limsup_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t}. \end{aligned}$$

按注 2.1.3, 对于任给的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $\|x' - x\| + t < \delta$ 时有

$$\begin{aligned} & (f_1 + f_2)^0(x; d) - \varepsilon \\ & \leq \frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t} \\ &= \frac{f_1(x' + td) - f_1(x')}{t} + \frac{f_2(x' + td) - f_2(x')}{t}. \end{aligned}$$

令 $x' \rightarrow x, t \rightarrow 0^+$, 从上式知

$$(f_1 + f_2)^0(x; d) \leq f_1^0(x; d) + f_2^0(x; d) \quad \forall d \in \mathcal{B}.$$

设 $x^* \in \partial^0(f_1 + f_2)(x)$, 由定义 4.2.1 和上式得

$$f_1^0(x; d) + f_2^0(x; d) \geq \langle x^*, d \rangle \quad \forall d \in \mathcal{B}.$$

据此, 由定理 4.2.1 的 (2) 知, 对任意的 $d \in \mathcal{B}$ 和任意的 $x^* \in \partial^0(f_1 + f_2)(x)$, 有

$$\begin{aligned} & \max\{\langle x_1^*, d \rangle \mid x_1^* \in \partial^0 f_1(x)\} + \max\{\langle x_2^*, d \rangle \mid x_2^* \in \partial^0 f_2(x)\} \\ & \geq \langle x^*, d \rangle, \end{aligned}$$

即

$$\begin{aligned} & \max\{\langle x_1^* + x_2^*, d \rangle \mid x_1^* + x_2^* \in \partial^0 f_1(x) + \partial^0 f_2(x)\} \\ & \geq \langle x^*, d \rangle \quad \forall d \in \mathcal{B}, \forall x^* \in \partial^0(f_1 + f_2)(x). \end{aligned} \tag{4.3.1}$$

现假设 $x^* \notin \partial^0 f_1(x) + \partial^0 f_2(x)$, 由定理 4.2.1 的(1)和定理 4.2.2 知 $\partial^0 f_1(x) + \partial^0 f_2(x)$ 是闭凸集, 因此单点集 $\{x^*\}$ 与凸集 $\partial^0 f_1(x) + \partial^0 f_2(x)$ 是可严格分离的. 根据定理 1.3.7 的(2), 存在 $\bar{d} \in \mathscr{D} \setminus \{0\}$, 使得

$$\langle x_1^* + x_2^*, \bar{d} \rangle < \langle x^*, \bar{d} \rangle$$

$$\forall x_1^* + x_2^* \in \partial^0 f_1(x) + \partial^0 f_2(x),$$

它与(4.3.1)相矛盾, 于是定理得证. \square

定理 4.3.3 设 f_1 和 f_2 与定理 4.3.2 的条件相同. 若 f_1 或 f_2 在 $x \in \mathscr{D}$ 处是严格可微的, 则

$$\partial^0(f_1 + f_2)(x) = \partial^0 f_1(x) + \partial^0 f_2(x).$$

证明 不妨设 f_1 在 x 处是严格可微的, 则对 $x' \in \mathscr{D}$ 和任给的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $\|x' - x\| + t < \delta$ 时有

$$\begin{aligned} \frac{f_1(x' + td) - f_1(x')}{t} - \varepsilon &\leq f'_1(x; d) \\ &\leq \frac{f_1(x' + td) - f_1(x')}{t} + \varepsilon. \end{aligned}$$

据此, 我们有

$$\begin{aligned} &\frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t} - \varepsilon \\ &\leq f'_1(x; d) + \frac{f_2(x' + td) - f_2(x')}{t} \\ &\leq \frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t} + \varepsilon. \end{aligned}$$

令 $x' \rightarrow x$, $t \rightarrow 0^+$, 得到

$$\begin{aligned} (f_1 + f_2)^0(x; d) - \varepsilon &\leq f'_1(x; d) + f_2^0(x; d) \\ &\leq (f_1 + f_2)^0(x; d) + \varepsilon. \end{aligned}$$

由 ϵ 的任意性, 从上式有

$$(f_1 + f_2)^0(x; d) = f_1^0(x; d) + f_2^0(x; d).$$

类似于定理 4.3.2 的证明可得到结论. \square

定理 4.3.4 设 f_1 和 f_2 同定理 4.3.2 的条件. 若 f_1 和 f_2 在 $x \in \mathscr{B}$ 处是正则的, 则

$$\partial^0(f_1 + f_2)(x) = \partial^0 f_1(x) + \partial^0 f_2(x).$$

证明 由定理 4.2.12 的(3), 有

$$(f_1 + f_2)^0(x; d) = f_1^0(x; d) + f_2^0(x; d),$$

再由定理 4.2.11 的(1)即可得证. \square

下面是中值定理.

定理 4.3.5 设 $x, y \in \mathscr{B}$, $S \subset \mathscr{B}$ 是开集, $[x, y] \subset S$. 若 f 在 S 上是 Lipschitz 的, 则存在点 $\xi \in (x, y)$, 使得

$$f(y) - f(x) \in \langle \partial^0 f(\xi), y - x \rangle.$$

证明 设 $\lambda \in [0, 1]$, 记 $x_\lambda = x + \lambda(y - x)$. 作函数 $g: [0, 1] \rightarrow R$, $g(\lambda) = f(x_\lambda)$, 显然 g 在 $(0, 1)$ 上是 Lipschitz 的. 令 $\delta = \pm 1$, 我们有

$$\begin{aligned} g^0(\lambda; \delta) &= \limsup_{\substack{\lambda' \rightarrow \lambda \\ t \rightarrow 0^+}} \frac{g(\lambda' + t\delta) - g(\lambda')}{t} \\ &= \limsup_{\substack{\lambda' \rightarrow \lambda \\ t \rightarrow 0^+}} \frac{f(x + (\lambda' + t\delta)(y - x)) - f(x + \lambda'(y - x))}{t} \\ &\leq \limsup_{\substack{\lambda' \rightarrow \lambda \\ t \rightarrow 0^+}} \frac{f(y' + t\delta(y - x)) - f(y')}{t} \\ &= f^0(x_\lambda; \delta(y - x)) \\ &= \max \langle \partial^0 f(x_\lambda), \delta(y - x) \rangle. \end{aligned}$$

于是,由定理 4.2.11 的(1)得 $\partial^0 g(\lambda) \subset \langle \partial^0 f(x_1), y - x \rangle$. 现考虑函数

$$\theta: [0, 1] \rightarrow R, \theta(\lambda) = g(\lambda) + \lambda[f(x) - f(y)],$$

则有 $\theta(0) = \theta(1) = f(x)$. 因为 θ 在 $(0, 1)$ 上是连续的,故它在 $(0, 1)$ 上可达到局部极小值. 由此,据定理 4.2.3 得 $0 \in \partial^0 \theta(\lambda)$, 再由定理 4.3.1 和定理 4.3.2 有

$$0 \in \langle \partial^0 f(x_1), y - x \rangle + f(x) - f(y).$$

取 $\xi = x_1$, 即得定理的结论. \square

以下介绍复合函数求 G -次微分的链式法则.

定理 4.3.6 设 $h: \mathcal{B} \rightarrow R^m$ 是向量函数, $h(x) = (h_1(x), \dots, h_m(x))^T$, $g: R^m \rightarrow R$ 和 $f: \mathcal{B} \rightarrow R$ 是实值函数, $f(x) = g(h(x))$. 若 $h_i (i = 1, \dots, m)$ 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, g 在 $h(x)$ 处是 Lipschitz 的, 则

$$\partial^0 f(x) \subset \overline{\text{co}} \left\{ \sum_{i=1}^m \lambda_i x_i^* \mid x_i^* \in \partial^0 h_i(x), i = 1, \dots, m; \right.$$

$$\left. \lambda \in \partial^0 g(h(x)), \lambda = (\lambda_1, \dots, \lambda_m)^T \right\}.$$

此外,使上式为等式,只要以下条件之一成立:

(1) g 在 $h(x)$ 处是正则的, $h_i (i = 1, \dots, m)$ 在 x 处是正则的, $\partial^0 g(h(x))$ 的元素的每一个分量是非负的(这时 f 在 x 处是正则的).

(2) g 在 $h(x)$ 处是严格可微的.

(3) g 在 $h(x)$ 处是正则的和 h 在 x 处是严格可微的(这时 f 在 x 处是正则的).

证明 记

$$S = \overline{\text{co}} \left\{ \sum_{i=1}^m \lambda_i x_i^* \mid x_i^* \in \partial^0 h_i(x), i = 1, \dots, m; \right.$$

$$\lambda \in \partial^0 g(\mathbf{h}(\mathbf{x})), \lambda = (\lambda_1, \dots, \lambda_m)^T \Big\},$$

设

$$\sigma_S(\mathbf{d}) = \max \left\{ \sum_{i=1}^m \lambda_i \langle \mathbf{x}_i^*, \mathbf{d} \rangle \mid \mathbf{x}_i^* \in \partial^0 h_i(\mathbf{x}), \right. \\ \left. i = 1, \dots, m; \lambda \in \partial^0 g(\mathbf{h}(\mathbf{x})) \right\}.$$

易知 σ_S 是 S 的支撑函数, 由定理 4.2.11, 只要证明 $\sigma_S(\mathbf{d}) \geq f^0(\mathbf{x}; \mathbf{d})$, 就有 $\partial^0 f(\mathbf{x}) \subset S$. 设 $\varepsilon > 0$, 作函数 $\sigma_\varepsilon: \mathcal{B} \rightarrow R$,

$$\sigma_\varepsilon(\mathbf{d}) = \max \left\{ \sum_{i=1}^m \lambda_i \langle \mathbf{x}_i^*, \mathbf{d} \rangle \mid \mathbf{x}_i^* \in \partial^0 h_i(\mathbf{x}_i), i = 1, \dots, m; \right. \\ \left. \lambda \in \partial^0 g(\boldsymbol{\xi}), \mathbf{x}_i \in \mathbf{x} + \varepsilon B, \boldsymbol{\xi} \in \mathbf{h}(\mathbf{x}) + \varepsilon B \right\}.$$

由定义 4.1.2, 对任给的 ε , 可以找到 \mathbf{x} 附近的点 \mathbf{x}' , 0 附近的 $t > 0$, 使得

$$f^0(\mathbf{x}; \mathbf{d}) \leq \frac{f(\mathbf{x}' + t\mathbf{d}) - f(\mathbf{x}')}{t} + \varepsilon. \quad (4.3.2)$$

取 \mathbf{x}' , $\mathbf{x}' + t\mathbf{d} \in \mathbf{x} + \varepsilon B$, 以及 $\mathbf{h}(\mathbf{x}')$, $\mathbf{h}(\mathbf{x}' + t\mathbf{d}) \in \mathbf{h}(\mathbf{x}) + \varepsilon B'$, 其中 $B' \subset R^m$ 是开单位球. 由定理 4.3.5 得

$$f(\mathbf{x}' + t\mathbf{d}) - f(\mathbf{x}') = g(\mathbf{h}(\mathbf{x}' + t\mathbf{d})) - g(\mathbf{h}(\mathbf{x}')) \\ = \sum_{i=1}^m \lambda_i [h_i(\mathbf{x}' + t\mathbf{d}) - h_i(\mathbf{x}')],$$

其中 $(\lambda_1, \dots, \lambda_m)^T \in \partial^0 g(\boldsymbol{\xi})$, $\boldsymbol{\xi}$ 是线段 $[\mathbf{h}(\mathbf{x}'), \mathbf{h}(\mathbf{x}' + t\mathbf{d})]$ 中的点, 从而 $\boldsymbol{\xi} \in \mathbf{h}(\mathbf{x}) + \varepsilon B'$. 再利用定理 4.3.5 可得

$$f(\mathbf{x}' + t\mathbf{d}) - f(\mathbf{x}') = \sum_{i=1}^m \lambda_i \langle \mathbf{x}_i^*, t\mathbf{d} \rangle,$$

其中 $\mathbf{x}_i^* \in \partial^0 h_i(\mathbf{x}_i)$, \mathbf{x}_i 是线段 $[\mathbf{x}', \mathbf{x}' + t\mathbf{d}]$ 中的点. 因而 $\mathbf{x}_i \in \mathbf{x} + \varepsilon B$. 将上式代入 (4.3.2), 有

$$f^0(x; d) \leq \sigma_\varepsilon(d) + \varepsilon, \quad (4.3.3)$$

下面证明 $\varepsilon \rightarrow 0^+$ 时, $\sigma_\varepsilon(d) \rightarrow \sigma_S(d)$. 设任给 $\delta > 0$, γ 是 $h_i (i = 1, \dots, m)$ 的 Lipschitz 常数, 不难选择 ε , 使得对每一 $h_i (i = 1, \dots, m)$, 它在 $x + \varepsilon B$ 上是 γ 秩 Lipschitz 的. 对每一 $x_i \in x + \varepsilon B$, 我们有

$$h_i^0(x_i; \pm d) \leq h_i^0(x; \pm d) + \frac{\delta}{\gamma}, \quad i = 1, \dots, m.$$

将上式两端同乘一个 $|\lambda_i|$, 其中 $(\lambda_1, \dots, \lambda_m)^T \in \partial^0 g(\xi)$, 注意 $|\lambda_i| \leq \gamma$, 则有

$$h_i^0(x_i; \lambda_i d) \leq h_i^0(x; \lambda_i d) + \delta, \quad i = 1, \dots, m. \quad (4.3.4)$$

根据定理 4.2.2 的(4), 也可选择 ε 使得

$$\partial^0 g(h(x)) + \varepsilon B \subset \partial^0 g(h(x)) + \delta B'.$$

利用(4.3.4)和上式, 有

$$\begin{aligned} \sigma_\varepsilon(d) &\leq \max \left\{ \sum_{i=1}^m \max [\lambda_i \langle x_i^*, d \rangle \mid x_i^* \in \partial^0 h_i(x_i), x_i \in x + \varepsilon B] \right. \\ &\quad \left. \mid (\lambda_1, \dots, \lambda_m)^T \in \partial^0 g(h(x)) + \varepsilon B' \right\} \\ &\leq \max \left\{ \sum_{i=1}^m (h_i^0(x, \lambda_i d) + \delta) \right. \\ &\quad \left. \mid (\lambda_1, \dots, \lambda_m)^T \in \partial^0 g(h(x)) + \delta B' \right\} \\ &\leq \max \left\{ \sum_{i=1}^m \max [\lambda_i \langle x_i^*, d \rangle \mid x_i^* \in \partial^0 h_i(x)] \right. \\ &\quad \left. \mid (\lambda_1, \dots, \lambda_m)^T \in \partial^0 g(h(x)) + \varepsilon B' \right\} + m\delta \\ &\leq \sigma_S(d) + n\delta\gamma \|d\| + m\delta. \end{aligned}$$

由 δ 的任意性和 $\sigma_S(d) \leq \sigma_\varepsilon(d)$, 可得 $\lim_{\varepsilon \rightarrow 0^+} \sigma_\varepsilon(d) = \sigma_S(d)$. 根据

(4.3.3)得到 $f^0(x; d) \leq \sigma_S(d)$, 因此 $\partial^0 f(x) \subset S$.

以下证明等式成立的情形.

(1) 因为已知 $\partial^0 g(h(x))$ 的元素 $\lambda = (\lambda_1, \dots, \lambda_m)^T$ 的每一分量 $\lambda_i (i = 1, \dots, m)$ 为非负, 我们有

$$\begin{aligned} \sigma_S(d) &= \max \left\{ \sum_{i=1}^m \max [\lambda_i \langle x_i^*, d \rangle \mid x_i^* \in \partial^0 h_i(x)] \right. \\ &\quad \left. \mid (\lambda_1, \dots, \lambda_m)^T \in \partial^0 g(h(x)) \right\} \\ &= \max \left\{ \sum_{i=1}^m \lambda_i w_i \mid \lambda \in \partial^0 g(h(x)) \right\}, \end{aligned}$$

其中 $w_i = \max [\langle x_i^*, d \rangle \mid x_i^* \in \partial^0 h_i(x)] (i = 1, \dots, m)$.

令 $w = (w_1, \dots, w_m)^T$, 从上式有

$$\sigma_S(d) = g'(h(x); w) = \lim_{t \rightarrow 0^+} \frac{g(h(x) + tw) - g(h(x))}{t},$$

或

$$\begin{aligned} \sigma_S(d) &= \lim_{t \rightarrow 0^+} \frac{g(h(x) + td) - g(h(x))}{t} \\ &\quad - \frac{g(h(x) + tw) - g(h(x + td))}{t}. \end{aligned} \quad (4.3.5)$$

注意 g 在 $h(x)$ 处是 Lipschitz 的, 故有

$$\begin{aligned} &\left| \frac{g(h(x) + tw) - g(h(x + td))}{t} \right| \\ &\leq \gamma \left\| w - \frac{h(x + td) - h(x)}{t} \right\|. \end{aligned}$$

当 $t \rightarrow 0^+$ 时, 上式右端趋于 0. 因此, 从 (4.3.5) 注意到 $f(x) = g(h(x))$, 有 $\sigma_S(d) = f'(x; d)$. 前面已经证明 $f^0(x; d) \leq \sigma_S(d)$, 由于 $f'(x; d) \leq f^0(x; d)$, 故得 $f^0(x; d) = f'(x; d)$. 由

此, $f^0(x; d)$ 是 S 的支撑函数, 利用定理 4.2.11, 便得到 $\partial^0 f(x) = S$.

(2) 由已设 g 在 $h(x)$ 处是严格可微的, 不妨设 $\beta = D_S g(h(x)) \geq 0$, 则有

$$\sigma_S(d) = \beta h^0(x; d) = \lim_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{\beta(h(x' + td) - h(x'))}{t}. \quad (4.3.6)$$

另外, 由定义 4.2.3 得

$$\beta = \lim_{\substack{h(x' + td) \rightarrow h(x') \\ x' \rightarrow x \\ t \rightarrow 0^+}} \frac{g(h(x' + td)) - g(h(x'))}{h(x' + td) - h(x')}.$$

从上式和 (4.3.6) 有

$$\sigma_S(d) = \lim_{\substack{x' \rightarrow x \\ t \rightarrow 0^+}} \frac{g(h(x' + td)) - g(h(x'))}{t} = f^0(x; d),$$

再根据定理 4.2.11, 即得 $\partial^0 f(x) = S$.

(3) 与 (1) 的证明类似. \square

定理 4.3.7 设 \mathcal{B} 是 Banach 空间, $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是映射, $g: \mathcal{D} \rightarrow R$ 是实值函数, $f = g \circ \varphi$. 若 φ 在 $x \in B$ 处是严格可微的, g 在 $\varphi(x)$ 处是 Lipschitz 的, 则 f 在 x 处是 Lipschitz 的, 并且有

$$\partial^0 f(x) \subset \partial^0 g(\varphi(x)) \circ D_S \varphi(x). \quad (4.3.7)$$

再若 g 在 $\varphi(x)$ 处是正则的, 则 f 在 x 处是正则的, 并且 (4.3.7) 成为等式. 注意 (4.3.7) 等价于: 若 $x^* \in \partial^0 f(x)$, 则存在 $y^* \in \partial^0 g(\varphi(x))$ 和 $D_S \varphi(x)$, 使得对任意的 $d \in \mathcal{B}$ 有

$$\langle x^*, d \rangle = \langle y^*, \langle D_S \varphi(x), d \rangle \rangle.$$

证明 根据定理 4.2.4, φ 在 x 处是严格可微的, 则 φ 在 x 处

是 Lipschitz 的. 于是, 存在 $\delta_1 > 0$, 使得对任意的 $x^1, x^2 \in x + \delta_1 B_{\mathcal{X}}$ 和某 $\gamma_1 > 0$ 有

$$\|\varphi(x^1) - \varphi(x^2)\|_{\mathcal{Z}} \leq \gamma_1 \|x^1 - x^2\|_{\mathcal{X}}. \quad (4.3.8)$$

由于 g 在 $\varphi(x)$ 处是 Lipschitz 的, 则存在 $\delta_2 > 0, \gamma_2 > 0$ 使得对任意的 $u^1, u^2 \in \varphi(x) + \delta_2 B_{\mathcal{Z}}$, 有

$$|g(u^1) - g(u^2)| \leq \gamma_2 \|u^1 - u^2\|_{\mathcal{Z}}. \quad (4.3.9)$$

因为 φ 在 x 处是连续的, 选择 $\delta > 0$, 使得当 $\delta \leq \delta_1$ 时, 对任意的 $x^1, x^2 \in x + \delta B_{\mathcal{X}}$ 有 $\varphi(x^1), \varphi(x^2) \in \varphi(x) + \delta_2 B_{\mathcal{Z}}$. 由 (4.3.8) 和 (4.3.9) 我们有

$$|f(x^1) - f(x^2)| \leq \gamma_1 \gamma_2 \|x^1 - x^2\|_{\mathcal{X}},$$

故 f 在 x 处是 Lipschitz 的.

设 $S = \partial^0 g(\varphi(x)) \circ D_S \varphi(x)$, 记 $\varphi' = D_S \varphi(x)$, 作

$$g^0(\varphi(x); \varphi'(d)) = \max \{ \langle y^*, \varphi'(d) \rangle \mid y^* \in \partial^0 g(\varphi(x)) \}.$$

因为 S 是凸的弱紧集, $g^0(\varphi(x); \varphi'(d))$ 是 S 的支撑函数, 故有

$$\begin{aligned} \sigma_S(d) &= \max \{ \langle x^*, d \rangle \mid x^* \in S \} \\ &= g^0(\varphi(x); \varphi'(d)). \end{aligned}$$

现在只要证明 $f^0(x; d) \leq g^0(\varphi(x); \varphi'(d))$, 根据定理 4.2.11 即可得 $\partial^0 f(x) \subset S$. 为此对任给的 $\varepsilon > 0$, 选择

$$\begin{aligned} x', x' + td &\in x + \varepsilon B_{\mathcal{X}}, \\ \varphi(x'), \varphi(x' + td) &\in \varphi(x) + \varepsilon B_{\mathcal{Z}}, \end{aligned}$$

由 $f^0(x; d)$ 的定义, 使得

$$f^0(x; d) \leq \frac{f(x' + td) - f(x')}{t} + \varepsilon. \quad (4.3.10)$$

另外,由定理 4.3.5 有

$$\begin{aligned} f(x' + td) - f(x') &= g(\varphi(x' + td)) - g(\varphi(x')) \\ &= \langle y^*, \varphi(x' + td) - \varphi(x') \rangle, \end{aligned}$$

其中 $y^* \in \partial^0 g(\xi)$, $\xi \in [\varphi(x'), \varphi(x' + td)]$, 因此 $\xi \in \varphi(x) + \varepsilon B_{\mathcal{L}}$. 令 $w = \varphi(x' + td) - \varphi(x')$, 由定理 4.1.2 知 $g^0(\varphi(x); w)$ 在 $(\varphi(x), w)$ 处是上半连续的. 选择适当的 ε , 使得 $\xi \in \varphi(x) + \varepsilon B_{\mathcal{L}}$, 有

$$g^0(\xi; w) \leq g^0(\varphi(x); w) + \varepsilon t.$$

由 (4.3.10) 知

$$\begin{aligned} f^0(x; d) &\leq \left\langle y^*, \frac{w}{t} \right\rangle + \varepsilon \\ &\leq \max \left\{ \left\langle y^*, \frac{w}{t} \right\rangle \mid y^* \in \partial^0 g(\varphi(x)) \right\} + \varepsilon, \end{aligned}$$

或

$$f^0(x; d) \leq g^0\left(\xi; \frac{w}{t}\right) + \varepsilon \leq g^0\left(\varphi(x); \frac{w}{t}\right) + 2\varepsilon.$$

上式即为

$$\begin{aligned} f^0(x; d) &\leq \max \left\{ \left\langle y^*, \frac{\varphi(x' + td) - \varphi(x')}{t} \right\rangle \mid y^* \in \partial^0 g(\varphi(x)) \right\} \\ &\quad + 2\varepsilon. \end{aligned}$$

令 $x' \rightarrow x$, $t \rightarrow 0^+$, 得

$$f^0(x; d) \leq \max \{ \langle y^*, \varphi'(d) \rangle \mid y^* \in \partial^0 g(\varphi(x)) \} + 2\varepsilon.$$

因为 ε 可任意小, 故得

$$f^0(x; d) \leq g^0(\varphi(x); \varphi'(d)). \quad (4.3.11)$$

若 g 在 $\varphi(x)$ 处是正则的, 则因为

$$\frac{[\varphi(x+td) - \varphi(x) - t\varphi'(d)]}{t} \rightarrow 0 \quad (t \rightarrow 0),$$

我们有

$$\begin{aligned} \lim_{t \rightarrow 0^+} \frac{g(\varphi(x) + t\varphi'(d)) - g(\varphi(x))}{t} &= \lim_{t \rightarrow 0^+} \frac{g(\varphi(x+td)) - g(\varphi(x))}{t} \\ &= \lim_{t \rightarrow 0^+} \frac{f(x+td) - f(x)}{t} \\ &= f'(x; d) \leq f^0(x; d). \end{aligned}$$

从上式和(4.3.11)得到 $f^0(x; d) = g^0(\varphi(x); \varphi'(d))$, 故 $\partial^0 f(x) = S$. \square

推论 4.3.8 设 $f_i: \mathcal{B} \rightarrow R$ ($i = 1, \dots, m$) 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, $f: \mathcal{B} \rightarrow R, f(x) = \max\{f_i(x) | i = 1, \dots, m\}$,

$$I(x) = \{i \in \{1, \dots, m\} | f_i(x) = f(x)\},$$

则

$$\partial^0 f(x) \subset \text{co}\{\partial^0 f_i(x) | i \in I(x)\}. \quad (4.3.12)$$

再若当 $i \in I(x)$ 时 f_i 在 x 处是正则的, 则(4.3.12)成为等式, 并且 f 在 x 处是正则的.

证明 定义函数 $g: R^m \rightarrow R$,

$$g(u_1, \dots, u_m) = \max\{u_i | i = 1, \dots, m\},$$

以及 $h: \mathcal{B} \rightarrow R^m, h(x) = (f_1(x), \dots, f_m(x))^T$, 则 $f(x) = g(h(x))$. 据定理 4.3.6 和例 4.2.2 即得结论. \square

下面阐述乘法规则和除法规则.

定理 4.3.9 设 $f_1, f_2: \mathcal{B} \rightarrow R$ 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, 则 $f_1 \cdot f_2$ 在 x 处是 Lipschitz 的, 并且

$$\partial^0(f_1 \cdot f_2)(x) \subset f_2(x)\partial^0 f_1(x) + f_1(x)\partial^0 f_2(x). \quad (4.3.13)$$

再若 $f_1(x) \geq 0$, $f_2(x) \geq 0$, 并且 f_1 和 f_2 均在 x 处是正则的, 则 $f_1 \cdot f_2$ 在 x 处是正则的, 并且 (4.3.13) 成为等式.

证明 作函数 $g: R^2 \rightarrow R$, $g(u_1, u_2) = u_1 u_2$, 以及 $h: \mathcal{B} \rightarrow R^2$, $h(x) = (f_1(x), f_2(x))^T$, 则 $(f_1 \cdot f_2)(x) = g(h(x))$. 根据定理 4.3.6 可以推得结论. \square

定理 4.3.10 设 $f_1, f_2: \mathcal{B} \rightarrow R$ 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, $f_2(x) \neq 0$, 则 f_1/f_2 在 x 处是 Lipschitz 的, 并且有

$$\partial^0(f_1/f_2)(x) \subset \frac{f_2(x)\partial^0 f_1(x) - f_1(x)\partial^0 f_2(x)}{f_2^2(x)}. \quad (4.3.14)$$

再若 $f_1(x) \geq 0$, $f_2(x) \geq 0$, 并且 f_1 和 $-f_2$ 均在 x 处是正则的, 则 f_1/f_2 在 x 处是正则的, 并且 (4.3.14) 成为等式.

证明 由定理 4.3.6 可以推得定理成立. \square

最后, 我们介绍 G -偏次微分的概念, 并且给出两个结论.

设 \mathcal{B}_1 和 \mathcal{B}_2 是 Banach 空间, $\mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2$, $f: \mathcal{B} \rightarrow R$, $x_1 \in \mathcal{B}_1$, $x_2 \in \mathcal{B}_2$, $f(x_1, x_2)$ 在 (x_1, x_2) 处是 Lipschitz 的. 固定 x_2 , 记 $f_1^0(x_1, x_2; d^1)$ 表示 f 在 x_1 处沿方向 $d^1 \in \mathcal{B}_1$ 的 G -方向导数, 则称 $f(\cdot, x_2)$ 在 x_1 处的 G -次微分 $\partial_1^0 f(x_1, x_2)$ 为 f 在 x_1 处的 G -偏次微分. 固定 x_1 , $f_2^0(x_1, x_2; d^2)$ 表示 f 在 x_2 处沿方向 $d^2 \in \mathcal{B}_2$ 的 G -方向导数, 则称 $f(x_1, \cdot)$ 在 x_2 处的 G -次微分 $\partial_2^0 f(x_1, x_2)$ 为 f 在 x_2 处的 G -偏次微分.

定理 4.3.11 设 f 在 $(x_1, x_2) \in \mathcal{B}_1 \times \mathcal{B}_2$ 处是正则的, 则

$$\partial^0 f(x_1, x_2) \subset \partial_1^0 f(x_1, x_2) \times \partial_2^0 f(x_1, x_2).$$

证明 设 $x^* = (x_1^*, x_2^*) \in \partial^0 f(x_1, x_2)$, 由定理 4.2.1 的 (2) 有

$$\begin{aligned} & f^0(x_1, x_2; (d, \theta)) \\ &= \max \{ \langle (x_1^*, x_2^*), (d, \theta) \rangle \mid x^* \in \partial^0 f(x) \} \end{aligned}$$

$$\geq \langle x_1^*, d \rangle \quad \forall d \in \mathcal{B}\}.$$

因为 f 在 x 处是正则的, 故有

$$\begin{aligned} f'_1(x_1, x_2; d) &= f'(x_1, x_2; (d, \theta)) \\ &= f^0(x_1, x_2; (d, \theta)), \end{aligned}$$

以及

$$f'_1(x_1, x_2; d) \geq \langle x_1^*, d \rangle \quad \forall d \in \mathcal{B}.$$

由此, 我们有 $x_1^* \in \partial_1^0 f(x_1, x_2)$. 同理, 可得 $x_2^* \in \partial_2^0 f(x_1, x_2)$. \square

定理 4.3.12 设 \mathcal{B}_1^* 和 \mathcal{B}_2^* 分别是 \mathcal{B}_1 和 \mathcal{B}_2 的对偶空间. 记

$$\begin{aligned} \pi_1 \partial^0 f(x_1, x_2) &= \{x_1^* \in \mathcal{B}_1^* \mid (x_1^*, x_2^*) \in \partial^0 f(x_1, x_2), x_2^* \in \mathcal{B}_2^*\}, \\ \pi_2 \partial^0 f(x_1, x_2) &= \{x_2^* \in \mathcal{B}_2^* \mid (x_1^*, x_2^*) \in \partial^0 f(x_1, x_2), x_1^* \in \mathcal{B}_1^*\}, \end{aligned}$$

则

$$\partial_1^0 f(x_1, x_2) \subset \pi_1 \partial^0 f(x_1, x_2),$$

$$\partial_2^0 f(x_1, x_2) \subset \pi_2 \partial^0 f(x_1, x_2).$$

证明 固定 x_2 , 作函数 $f_1: \mathcal{B}_1 \rightarrow R$, $f_1(x) = f(x_1, x_2)$, 以及映射 $F: \mathcal{B}_1 \rightarrow \mathcal{B}_1 \times \mathcal{B}_2$, $F(x) = (x_1, x_2)$, 则 $\langle D_S F(x), d \rangle = \langle d, \theta \rangle$. 因为 $f_1 = f \circ F$, 利用定理 4.3.7 即得

$$\partial_1^0 f(x_1, x_2) \subset \pi_1 \partial^0 f(x_1, x_2).$$

同理, 可证第 2 个结论成立. \square

例 4.3.1 已知平面上的 $N+1$ 个点: $(0, 0)$, $(1, 1)$, $(2, 2)$, \dots , $(N-1, N-1)$ 和 $(N, 0)$, 考虑如何寻找一条直线来拟合这些点(图 4.3.1). 设直线方程为 $y = ax + \beta$, 问题归结为求以下的极小化问题:

$$\min f(\alpha, \beta) = \sum_{i=0}^{N-1} |\alpha i + \beta - i| + |\alpha N + \beta - 0|.$$

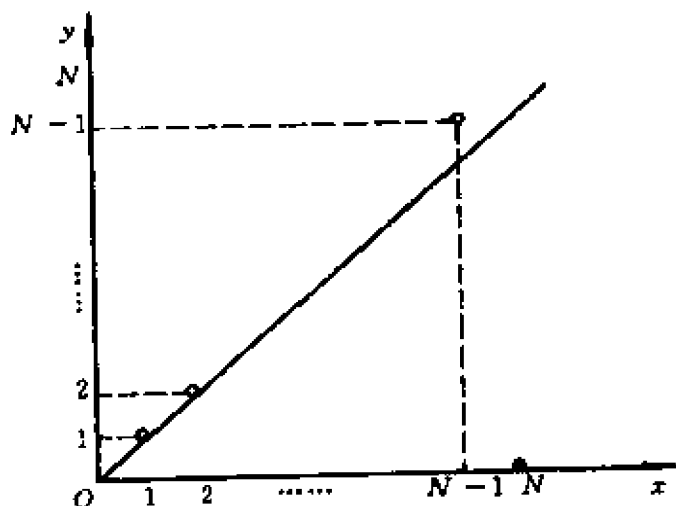


图 4.3.1

作函数 $f_{\alpha, c}: R^2 \rightarrow R$, $f_{\alpha, c}(\alpha, \beta) = |\alpha\alpha + \beta - c|$, 它是以下两函数的复合: $g: R \rightarrow R$, $h: R^2 \rightarrow R$,

$$g(u) = |u|, \quad h(\alpha, \beta) = \alpha\alpha + \beta - c.$$

注意, h 是严格可微的, g 是凸函数, 并且是正则的, 由

$$\partial^0 g(u) = \begin{cases} 1, & u > 0; \\ [-1, 1], & u = 0; \\ -1, & u < 0, \end{cases}$$

以及 $D_\delta h(\alpha, \beta) = (\alpha, 1)^T$, 利用定理 4.3.7 得

$$\partial^0 f_{\alpha, c}(\alpha, \beta) = \begin{cases} (\alpha, 1)^T, & \alpha\alpha + \beta - c > 0; \\ \{\gamma(\alpha, 1)^T \mid |\gamma| \leq 1\}, & \alpha\alpha + \beta - c = 0; \\ (-\alpha, -1)^T, & \alpha\alpha + \beta - c < 0. \end{cases} \quad (4.3.15)$$

由于 $f(\alpha, \beta) = \sum_{i=0}^{N-1} f_{i, i}(\alpha, \beta) + f_{N, 0}(\alpha, \beta)$, 根据定理 4.3.4 有

$$\partial^0 f(\alpha, \beta) = \sum_{i=0}^{N-1} \gamma_i(i, 1)^T + \gamma_N(N, 1)^T,$$

其中 $|\gamma_i| \leq 1$ ($i = 0, 1, \dots, N-1$). 对于直线 $y = x$ (即 $\alpha = 1, \beta = 0$), 则当 $a = N, c = 0$ 时, 有 $a\alpha + \beta - c = N > 0$, 由 (4.3.15) 知 $\gamma_N = 1$, 从而

$$\partial^0 f(1, 0) = \sum_{i=0}^{N-1} \gamma_i(i, 1)^T + \gamma_N(N, 1)^T. \quad (4.3.16)$$

由定理 4.2.3, $(1, 0)$ 是 $f(\alpha, \beta)$ 的极小点的必要条件是 $(0, 0)^T \in \partial^0 f(1, 0)$, 因为 f 是凸的, 根据定理 3.4.1, 它也是充分条件. 据此, 由 (4.3.16) 得知 $y = x$ 是问题的解的充要条件为

$$(0, 0)^T \in \sum_{i=0}^{N-1} \gamma_i(i, 1)^T + \gamma_N(N, 1)^T, \quad (4.3.17)$$

其中 $|\gamma_i| \leq 1$ ($i = 0, 1, \dots, N-1$). 容易看出, 当 $N \geq 3$ 时, 存在 γ_i ($i = 0, 1, \dots, N-1$) 满足 (4.3.17).

§ 4.4 几何特性

这一节, 利用集合的距离函数, 引进集合在一点处的 G -切锥和 G -法锥的概念. 由此, 来刻画 G -次微分的有关几何特性.

设 \mathcal{B} 是 Banach 空间, 集合 $S \subset \mathcal{B}$ 非空, 考虑 S 上的距离函数

$$d_S(x) = \inf_{y \in S} \|x - y\|, \quad x \in \mathcal{B}. \quad (4.4.1)$$

若 S 是闭的, 则 $x \in S$ 当且仅当 $d_S(x) = 0$.

定理 4.4.1 设集合 $S \subset \mathcal{B}$ 非空, 则距离函数 $d_S: \mathcal{B} \rightarrow R$ 在 \mathcal{B} 上是 Lipschitz 的, 即有

$$|d_S(x^1) - d_S(x^2)| \leq \|x^1 - x^2\| \quad \forall x^1, x^2 \in \mathcal{B}. \quad (4.4.2)$$

证明 由(4.4.1), 对于任给的 $\epsilon > 0$, 存在 $\bar{y} \in S$ 使得

$$d_S(x^2) \geq \|x^2 - \bar{y}\| - \epsilon.$$

于是, 我们有

$$\begin{aligned} d_S(x^1) &\leq \|x^1 - \bar{y}\| \leq \|x^1 - x^2\| + \|x^2 - \bar{y}\| \\ &\leq \|x^1 - x^2\| + d_S(x^2) + \epsilon. \end{aligned}$$

由 ϵ 的任意性得(4.4.2)成立. \square

设 \mathcal{B}^* 是 \mathcal{B} 的对偶空间.

定义 4.4.1 设集合 $S \subset \mathcal{B}$ 非空, 点 $x^0 \in \mathcal{B}$.

(1) 称集合

$$T_S^0(x^0) = \{d \in \mathcal{B} \mid d_S^0(x^0; d) = 0\} \quad (4.4.3)$$

是集合 S 在点 x^0 处的广义切锥, 或 G-切锥.

(2) 称集合

$$N_S^0(x^0) = \{x^* \in \mathcal{B}^* \mid \langle x^*, d \rangle \leq 0 \quad \forall d \in T_S^0(x^0)\} \quad (4.4.4)$$

是集合 S 在点 x^0 处的广义法锥, 或 G-法锥.

注 4.4.1 设 $x \in \mathcal{B}$, 由定义 4.4.1 易知 $T_S^0(x)$ 是闭凸集.

定理 4.4.2 设集合 $S \subset \mathcal{B}$ 非空, $x \in \mathcal{B}$, 则

$$N_S^0(x) = \text{cl} \bigcup_{\lambda \geq 0} \lambda \partial^0 d_S(x), \quad (4.4.5)$$

其中 cl 表示弱闭.

证明 由定义 4.4.1 的(1)和定理 4.2.1 的(2)可知, $d \in T_S^0(x)$ 当且仅当 $\langle x^*, d \rangle \leq 0$ 对任意的 $x^* \in \partial^0 d_S(x)$ 成立. 因为 $T_S^0(x)$ 是弱闭凸锥, 故(4.4.5)成立. \square

定理 4.4.3 设函数 $f: \mathcal{B} \rightarrow \mathbb{R}$ 在集合 $X \subset \mathcal{B}$ 上是 γ 秩 Lipschitz 的, $\bar{x} \in S \subset X$. 若 f 在 $\bar{x} \in S$ 处达到极小值, 则对任意的 $\tilde{\gamma} \geq \gamma$, 函数 $F(y) = f(y) + \tilde{\gamma} d_S(y)$ 在 $\bar{x} \in S$ 处达到极小值. 若

$\tilde{\gamma} > \gamma$, S 是闭的, 并且 $f(\bar{y}) = \min_{x \in X} F(x)$, 则 $\bar{y} \in S$.

证明 用反证法. 假设存在 $\bar{y} \in X$ 和 $\bar{\epsilon} > 0$, 使得

$$f(\bar{y}) + \tilde{\gamma} d_S(\bar{y}) < f(\bar{x}) - \tilde{\gamma} \bar{\epsilon}.$$

取 $x' \in S$ 使 $\|\bar{y} - x'\| \leq d_S(\bar{y}) + \bar{\epsilon}$, 由于 f 是 γ 秩 Lipschitz 的, 则有

$$\begin{aligned} f(x') &\leq f(\bar{y}) + \gamma \|\bar{y} - x'\| \\ &\leq f(\bar{y}) + \tilde{\gamma} (d_S(\bar{y}) + \bar{\epsilon}) < f(\bar{x}). \end{aligned}$$

上式与 f 在 $\bar{x} \in S$ 处达到极小值相矛盾. 现在设 $\tilde{\gamma} > \gamma$, 函数 $F(y) = f(y) + \tilde{\gamma} d_S(y)$ 在 X 上的点 \bar{y} 处达到极小值, 由前面所证, 可知有

$$f(\bar{y}) + \tilde{\gamma} d_S(\bar{y}) = f(\bar{x}) \leq f(\bar{x}) + \frac{\gamma + \tilde{\gamma}}{2} d_S(\bar{y}).$$

从上式得知 $d_S(\bar{y}) = 0$, 故 $\bar{y} \in S$. \square

推论 4.4.4 设函数 $f: \mathcal{B} \rightarrow R$ 在集合 $S \subset \mathcal{B}$ 上是 Lipschitz 的. 若 $f(\bar{x}) = \min_{x \in S} f(x)$, 则

$$0 \in \partial^0 f(\bar{x}) + N_S^0(\bar{x}).$$

证明 设 $U(\bar{x})$ 是 \bar{x} 的邻域, f 在 $U(\bar{x})$ 上是 γ 秩 Lipschitz 的. 不妨设 $S \subset U(\bar{x})$, 则由定理 4.4.3 知 $f(y) + \gamma d_S(y)$ 在 \bar{x} 处达到极小值. 因此, 有

$$\begin{aligned} 0 &\in \partial^0(f + \gamma d_S)(\bar{x}) \subset \partial^0 f(\bar{x}) + \gamma \partial^0 d_S(\bar{x}) \\ &\subset \partial^0 f(\bar{x}) + N_S^0(\bar{x}). \quad \square \end{aligned}$$

定理 4.4.5 设 $S \subset \mathcal{B}$ 是非空凸集, $x \in \mathcal{B}$, $N_S(x)$ 是 S 在点 x 处的法锥, 即

$$N_S(x) = \{x^* \in \mathcal{B}^* \mid \langle x^*, y - x \rangle \leq 0 \quad \forall y \in S\},$$

则 $N_S^0(x) = N_S(x)$.

证明 设 $x^* \in N_S(x)$, $f(y) = \langle x^*, x - y \rangle$, 我们有 $f(x) = \min_{y \in S} f(y)$. 于是, 根据推论 4.4.4 知 $0 \in -x^* + N_S^0(x)$, 故得 $x^* \in N_S^0(x)$.

反之, 设 $x^* \in N_S^0(x)$, 令 $x, x' \in \mathcal{B}$, $\lambda \in (0, 1)$. 对任给的 $\varepsilon > 0$, 选择 $x^1, x^2 \in S$ 使得

$$\|x^1 - x\| \leq d_S(x) + \varepsilon, \quad \|x^2 - x'\| \leq d_S(x') + \varepsilon,$$

则

$$\begin{aligned} d_S(\lambda x + (1 - \lambda)x') &\leq \|\lambda x^1 + (1 - \lambda)x^2 - \lambda x - (1 - \lambda)x'\| \\ &\leq \lambda d_S(x) + (1 - \lambda)d_S(x') + \varepsilon. \end{aligned}$$

因为 ε 是任意的, 故 d_S 是凸函数, 于是由定理 4.4.2 得 $x^* \in \partial^0 d_S(x)$. 利用定理 4.2.10, 知

$$\langle x^*, y - x \rangle \leq d_S(y) - d_S(x) \quad \forall y \in \mathcal{B}.$$

因为对任意的 $y \in S$, 有 $\langle x^*, y - x \rangle \leq 0$, 所以 $x^* \in N_S(x)$. \square

注 4.4.2 设 $S \subset \mathcal{B}$ 是非空凸集, $x \in \mathcal{B}$. 由定义 3.2.2, 定义 4.4.1 和定理 4.4.5 可知, 也有 $T_S^0(x) = T_S(x)$.

推论 4.4.6 设 $S \subset \mathcal{B}$ 是非空凸集, $x \in \mathcal{B}$, 则 $d \in T_S^0(x)$ 当且仅当 $d_S^0(x; d) = d'_S(x; d) = 0$.

证明 因为 d_S 是凸函数, 由定理 4.2.12 的(2)知它是正则的, 从而有 $d_S^0(x; d) = d'_S(x; d) = 0$. \square

定理 4.4.7 设集合 $S \subset \mathcal{B}$ 非空, $x \in \mathcal{B}$, 则 $d \in T_S^0(x)$ 当且仅当对任意的点列 $\{x^k\} \subset S$, $x^k \rightarrow x$, $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0$ ($k \rightarrow \infty$), 存在序列 $\{d^k\} \subset \mathcal{B}$, $d^k \rightarrow d$, 使得对任何 $k = 1, 2, \dots$, 有 $x^k + t_k d^k \in S$.

证明 设 $d \in T_S^0(x)$, $\{x^k\} \subset S$, $x^k \rightarrow x$, $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0$ ($k \rightarrow \infty$). 因为 $d_S^0(x; d) = 0$, d_S 是凸函数, 故有

$$\lim_{k \rightarrow \infty} \frac{d_S(x^k + t_k d) - d_S(x^k)}{t_k} = \lim_{k \rightarrow \infty} \frac{d_S(x^k + t_k d)}{t_k} = 0. \quad (4.4.6)$$

选择点 $y^k \in S$, 使得

$$\|x^k + t_k d - y^k\| \leq d_S(x^k + t_k d) + t_k/k. \quad (4.4.7)$$

设 $d^k = \frac{y^k - x^k}{t_k}$, 由(4.4.6)和(4.4.7), 令 $k \rightarrow \infty$, 则有 $d^k \rightarrow d$. 故对任何 $k = 1, 2, \dots$, 有 $x^k + t_k d^k \in S$.

反之, 设 $\{x^k\} \subset S$, $x_k \rightarrow x$, $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0$ ($k \rightarrow \infty$), 有

$$d_S^0(x; d) = \lim_{k \rightarrow \infty} \frac{d_S(x^k + t_k d) - d_S(x^k)}{t_k}. \quad (4.4.8)$$

选择 $y^k \in S$, 使得

$$\|y^k - x^k\| \leq d_S(x^k) + t_k/k. \quad (4.4.9)$$

在上式中令 $k \rightarrow \infty$, 有 $y^k \rightarrow x$, 故存在序列 $\{d^k\} \subset \mathcal{B}$, $d^k \rightarrow d$, 使得 $y^k + t_k d^k \in S$. 据定理 4.4.1, 我们有

$$d_S(x^k + t_k d) \leq d_S(y^k + t_k d^k) + \|x^k - y^k\| + t_k \|d^k - d\|.$$

由(4.4.9)得

$$d_S(x^k + t_k d) \leq d_S(x^k) + t_k (\|d^k - d\| + 1/k).$$

利用上式和(4.4.8), 令 $k \rightarrow \infty$, 有 $d_S^0(x; d) = 0$, 故 $d \in T_S^0(x)$. \square

推论 4.4.8 设 $\mathcal{B}_1, \mathcal{B}_2$ 是 Banach 空间, $S_1 \subset \mathcal{B}_1, S_2 \subset \mathcal{B}_2$, $x = (x_1, x_2) \in S_1 \times S_2$, 则

$$T_{S_1 \times S_2}^0(x) = T_{S_1}^0(x_1) \times T_{S_2}^0(x_2),$$

$$N_{S_1 \times S_2}^0(x) = N_{S_1}^0(x_1) \times N_{S_2}^0(x_2).$$

证明 由定理 4.4.7 容易推得. \square

定义 4.4.2 设集合 $S \subset \mathcal{B}$ 非空, 点 $x^0 \in S, d \in \mathcal{B}$. 若对任给的 $\epsilon > 0$, 存在 $t \in (0, \epsilon)$ 和 $p \in d + \epsilon B$ 使得 $x^0 + tp \in S$, 则由所有这样的 d 组成的集合称为 S 在点 x^0 处的相依锥, 记作

$Q_S(x^0)$. 若有 $T_S^0(x^0) = Q_S(x^0)$, 则称集合 S 在点 x^0 处是正则的.

注 4.4.3 设 $x \in S$, 据定理 4.4.7 易知有 $T_S^0(x) \subset Q_S(x)$, 但反之不成立.

定理 4.4.9 设函数 $f: \mathcal{B} \rightarrow R$ 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, $0 \notin \partial^0 f(x)$. 若 $S = \{y \in \mathcal{B} | f(y) \leq f(x)\}$, 则

$$\{d \in \mathcal{B} | f^0(x; d) \leq 0\} \subset T_S^0(x). \quad (4.4.10)$$

再若 f 在 x 处是正则的, 则集合 S 在 x 处是正则的, 并且 (4.4.10) 成为等式.

证明 设 $d \in \mathcal{B}$ 使 $f^0(x; d) < 0$, 则存在 $\bar{\epsilon}, \bar{\delta} > 0$, 使得对任意的 $x' \in x + \bar{\epsilon}B$ 和 $t \in (0, +\infty)$, 有

$$f(x' + td) - f(x') \leq -\bar{\delta}t.$$

设 $\{x^k\} \subset S$, $x^k \rightarrow x$, $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0$ ($k \rightarrow \infty$), 由 $f(x^k) \leq f(x)$, 故对任何 $k (= 1, 2, \dots)$, 有

$$f(x^k + t_k d) \leq f(x^k) - \bar{\delta}t_k \leq f(x) - \bar{\delta}t_k \leq f(x),$$

从而 $x^k + t_k d \in S$. 由此, 根据定理 4.4.7 得 $d \in T_S^0(x)$. 因为 $0 \notin \partial^0 f(x)$, 则存在 $d' \in S$, 使得 $f^0(x; d') < 0$, 即有 $d' \in T_S^0(x)$. 设 $d \in \{d \in \mathcal{B} | f^0(x; d) \leq 0\}$, 由定理 4.1.1 的(1)可知, 对任意的 $\epsilon > 0$ 有

$$f^0(x; d + \epsilon d') \leq f^0(x; d) + \epsilon f^0(x; d') < 0,$$

从而

$$d + \epsilon d' \in T_S^0(x).$$

由 $T_S^0(x)$ 是闭的和 ϵ 的任意性, 于是得 $d \in T_S^0(x)$.

现设 $d \in Q_S(x)$, 由定义 4.4.2 有

$$\liminf_{t \rightarrow 0^+} \frac{d_S(x + td)}{t} = 0,$$

即对任给的 $\epsilon > 0$, 存在收敛于 0 的正数列 $\{t_k\}$, 使得对充分大的 k

有 $d_S(\mathbf{x} + t_k \mathbf{d}) \leq \varepsilon t_k$, 并存在点列 $\{\mathbf{x}^k\} \subset S$ 使得

$$\|\mathbf{x} + t_k \mathbf{d} - \mathbf{x}^k\| \leq 2\varepsilon t_k.$$

根据 f 在 \mathbf{x} 处是 γ 秩 Lipschitz 的和 $f(\mathbf{x}^k) \leq f(\mathbf{x})$, 我们有

$$\frac{f(\mathbf{x} + t_k \mathbf{d}) - f(\mathbf{x})}{t_k} \leq \frac{f(\mathbf{x}^k) + 2\varepsilon \gamma t_k - f(\mathbf{x})}{t_k} \leq 2\varepsilon \gamma.$$

由于 f 在 \mathbf{x} 处是正则的, 在上式中令 $t_k \rightarrow 0$, $\varepsilon \rightarrow 0$ 可知 $f'(\mathbf{x}; \mathbf{d}) = f^0(\mathbf{x}; \mathbf{d}) \leq 0$, 因此有 $\{\mathbf{d} \in \mathcal{B} | f^0(\mathbf{x}; \mathbf{d}) \leq 0\} = T_S^0(\mathbf{x})$. \square

推论 4.4.10 设条件与定理 4.4.9 相同, 则

$$N_S^0(\mathbf{x}) \subset \bigcup_{\lambda \geq 0} \lambda \partial^0 f(\mathbf{x}).$$

再若 f 在 \mathbf{x} 处是正则的, 则上式成为等式.

证明 由定理 4.4.2 和定理 4.4.9 可直接推得. \square

推论 4.4.11 设 $f_i: \mathcal{B} \rightarrow R (i=1, 2, \dots, m)$ 是实值函数, 集合 $S = \{\mathbf{y} \in \mathcal{B} | f_1(\mathbf{y}) \leq 0, \dots, f_m(\mathbf{y}) \leq 0\}$, $\mathbf{x} \in \mathcal{B}$ 使 $f_i(\mathbf{x}) = 0 (i=1, \dots, m)$. 若 $f_i (i=1, \dots, m)$ 在 \mathbf{x} 处均是严格可微的, $D_S f_1(\mathbf{x}), \dots, D_S f_m(\mathbf{x})$ 是正线性独立的, 并且 S 在 \mathbf{x} 处是正则的, 则

$$N_S^0(\mathbf{x}) = \left\{ \sum_{i=1}^m \lambda_i D_S f_i(\mathbf{x}) | \lambda_i \geq 0, i=1, \dots, m \right\}.$$

证明 作函数 $f(\mathbf{y}) = \max\{f_i(\mathbf{y}) | i=1, \dots, m\}$. 根据定理 4.2.12 可知, f 在 \mathbf{x} 处是 Lipschitz 的和正则的. 由 $S = \{\mathbf{y} \in \mathcal{B} | f(\mathbf{y}) \leq 0\}$ 和 $f(\mathbf{x}) = 0$, 根据推论 4.4.10 即可推得结论. \square

下面介绍另一个切锥的概念, 并考虑它与 G -切锥之间的关系.

定义 4.4.3 设集合 $S \subset \mathcal{B}$ 非空, 点 $\mathbf{x}^0 \in S$, $\mathbf{d} \in \mathcal{B}$. 若对某个 $\varepsilon > 0$, 有 $\mathbf{p} \in \mathbf{d} + \varepsilon B$, $t \in (0, \varepsilon)$, 使得

$$\mathbf{x}' + t\mathbf{p} \in S \quad \forall \mathbf{x}' \in (\mathbf{x}^0 + \varepsilon B) \cap S,$$

则称 d 是集合 S 在 x^0 处的超切向量. S 在点 x^0 处的所有超切向量组成的集合称为 S 在点 x^0 处的超切锥, 或 H -切锥, 记作 $T_S^H(x^0)$.

注 4.4.4 设 $x \in \mathcal{B}$, 由定理 4.4.7 容易推得 $T_S^H(x) \subset T_S^0(x)$.

定理 4.4.12 设集合 $S \subset \mathcal{B}$ 非空, $x \in \mathcal{B}$. 若 $T_S^H(x) \neq \emptyset$, 则

$$T_S^H(x) = \text{int} T_S^0(x).$$

证明 首先证明下式成立:

$$T_S^H(x) + T_S^0(x) \subset T_S^H(x). \quad (4.4.11)$$

事实上, 设 $d^1 \in T_S^H(x)$, $d^2 \in T_S^0(x)$. 根据定义 4.4.3, 存在 $\varepsilon_1 > 0$ 使得

$$S \cap (x + \varepsilon_1 B) + t(d^1 + \varepsilon_1 B) \subset S \quad \forall t \in (0, \varepsilon_1). \quad (4.4.12)$$

再由定理 4.4.7, 存在 $\varepsilon_2 \leq 0$ 使得

$$S \cap (x + \varepsilon_2 B) + td^2 \subset S + t(\varepsilon_1/2)B, \quad t \in (0, \varepsilon_2). \quad (4.4.13)$$

取

$$\varepsilon < \min(\varepsilon_1/2, \varepsilon_1/(1 + \varepsilon_1 + \|d^2\|), \varepsilon_2),$$

$$d = y + t(d^1 + d^2 + \varepsilon u),$$

其中 $y \in S \cap (x + \varepsilon B)$, $u \in B$, $t \in (0, \varepsilon)$. 因为 $\varepsilon \leq \varepsilon_2$, 由 (4.4.13) 知, 对某个 $p \in B$ 有

$$y + td^2 - (1/2)t\varepsilon_1 p \in S.$$

另一方面, 有

$$\|x - y - td^2 + (1/2)t\varepsilon_1 p\| \leq \varepsilon + \varepsilon(\|d^2\| + \varepsilon_1) \leq \varepsilon_1.$$

从上面两式可得 $y + t(d^2 - (1/2)\varepsilon_1 p) \in S \cap (x + \varepsilon_1 B)$. 由

(4.4.12)有

$$y + t(d^2 - (1/2)\epsilon_1 p) + t(d^1 + \epsilon_1 B) \subset S,$$

从而

$$\begin{aligned} d &= y + t(d^2 - (1/2)\epsilon_1 p) + td^1 \\ &\quad + t(\epsilon u + (1/2)\epsilon_1 p) \subset S. \end{aligned}$$

注意 $\|\epsilon u + (1/2)\epsilon_1 p\| \leq \epsilon + (1/2)\epsilon_1 < \epsilon_1$, 根据定义 4.4.3 得 $d^1 + d^2 \in T_S^H(x)$, 因此(4.4.11)成立.

显然, $T_S^H(x)$ 是开集, 并且 $T_S^H(x) \subset T_S^0(x)$, 故只要证明 $\text{int}T_S^0(x) \subset T_S^H(x)$. 设 $d \in \text{int}T_S^0(x)$, 则对任意的 $p \in T_S^H(x)$, 存在 $\bar{\lambda} > 0$ 使得 $d - \bar{\lambda}p \in \text{int}T_S^0(x)$. 因为 $\bar{\lambda}p \in T_S^H(x)$, 利用(4.4.11), 便得 $d = \bar{\lambda}p + (d - \bar{\lambda}p) \in T_S^H(x)$. \square

推论 4.4.13 设集合 $S \subset \mathscr{B}$ 非空, $x \in S$. 若 $T_S^H(x) \neq \emptyset$, 则对任何的 $\{x_k\} \subset \mathscr{B}$, $\{x_k^*\} \subset \mathscr{B}^*$, $x_k^* \in N_S^0(x_k)$, $x_k^* \rightarrow x^* (k \rightarrow \infty)$, 有 $x^* \in N_S^0(x)$.

证明 设 $d \in T_S^H(x)$, 则存在 $\bar{\epsilon} > 0$ 使得

$$S \cap (x + \bar{\epsilon}B) + t(d + \bar{\epsilon}B) \subset S \quad \forall t \in (0, \epsilon).$$

由此, 对任意的 $y \in x + (\bar{\epsilon}/2)B$, 有

$$\begin{aligned} &S \cap \left(y + \frac{\bar{\epsilon}}{2}B\right) + t\left(d + \frac{\bar{\epsilon}}{2}B\right) \\ &\subset S \cap (x + \bar{\epsilon}B) + t(d + \bar{\epsilon}B) \\ &\subset S \quad \forall t \in \left(0, \frac{\bar{\epsilon}}{2}\right). \end{aligned}$$

根据定义 4.4.3, 对任意的 $y \in x + \frac{\bar{\epsilon}}{2}B$ 得 $d \in T_S^H(y)$, 因此对任意的 $y \in x + \frac{\bar{\epsilon}}{2}B$ 有 $d \in T_S^0(y)$. 由 $x_k \rightarrow x$, $x_k^* \rightarrow x^* (k \rightarrow \infty)$, 以及 $x_k^* \in N_S^0(x_k)$, 有

$$\langle x^*, d \rangle = \lim_{k \rightarrow \infty} \langle x_k^*, d \rangle \leqslant 0 \quad \forall d \in T_S^H(x).$$

于是, 由于定理 4.4.12 和 $T_S^0(x)$ 是闭的, 得到对任意的 $d \in T_S^0(x)$ 有 $\langle x^*, d \rangle \leqslant 0$, 从而 $x^* \in N_S^0(x)$. \square

从以下的讨论可知, 函数的上图象的 G -切锥和 G -法锥, 分别与它的 G -方向导数和 G -次梯度有着等价关系. 因此, 对于一般函数, 我们可以用 G -切锥和 G -法锥来引进导数和次梯度. 由定义 2.1.6 的 (1), 函数 $f: \mathcal{B} \rightarrow R$ 的上图象为

$$\text{epi} f = \{(x, \eta) \in \mathcal{B} \times R \mid \eta \geqslant f(x)\}.$$

定理 4.4.14 设函数 $f: \mathcal{B} \rightarrow R$ 在 $x \in \mathcal{B}$ 处是 Lipschitz 的.

(1) $(d, \eta) \in T_{\text{epi} f}^0(x, f(x))$ 当且仅当 $f^0(x; d) \leqslant \eta$.

(2) f 在 x 处是正则的当且仅当集合 $\text{epi} f$ 在 $(x, f(x))$ 处是正则的.

证明 (1) 设 $(d, \eta) \in T_{\text{epi} f}^0(x, f(x))$. 选取点列 $\{x^k\} \subset \mathcal{B}$, $x_k \rightarrow x$, $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0^+ (k \rightarrow \infty)$, 使得

$$\lim_{k \rightarrow \infty} \frac{f(x^k + t_k d) - f(x^k)}{t_k} = f^0(x; d).$$

因为 $(x^k, f(x^k)) \rightarrow (x, f(x))$, 由定理 4.4.7 知存在序列 $\{d^k, \eta_k\} \subset \mathcal{B} \times R$, $(d^k, \eta_k) \rightarrow (d, \eta)$, 使得 $(x^k, f(x^k)) + t_k(d^k, \eta_k) \in \text{epi} f$. 因此, 我们有

$$f(x^k) + t_k \eta_k \geqslant f(x^k + t_k d^k), \quad k = 1, 2, \dots,$$

即

$$\frac{f(x^k + t_k d) - f(x^k)}{t_k} \leqslant \eta_k, \quad k = 1, 2, \dots.$$

在上式中令 $k \rightarrow \infty$, 得 $f^0(x; d) \leqslant \eta$.

设序列 $\{(x^k, \eta_k)\} \subset \text{epi} f$, $x^k \rightarrow x$, $\eta_k \rightarrow f(x)$, 对任给的 $\delta > 0$, 取

$$s_k = \max \left\{ f^0(x; d) + \delta, \frac{f(x^k + t_k d) - f(x^k)}{t_k} \right\},$$

$$k = 1, 2, \dots,$$

我们有

$$\eta_k + t_k s_k \geq \eta_k + f(x^k + t_k d) - f(x^k), \quad k = 1, 2, \dots.$$

由于 $\eta_k \geq f(x^k)$, 故得

$$\eta_k + t_k s_k \geq f(x^k + t_k d), \quad k = 1, 2, \dots.$$

因为

$$\limsup_{k \rightarrow \infty} \frac{f(x^k + t_k d) - f(x^k)}{t_k} \leq f^\circ(x; d),$$

所以当 $k \rightarrow \infty$ 时有 $s_k \rightarrow f^\circ(x; d) + \delta$. 由此, 根据定理 4.4.7 得 $(d, f^\circ(x; d) + \delta) \in T_{\text{epi}f}^\circ(x, f(x))$, 也就是当 $f^\circ(x; d) \leq \eta$ 时有 $(d, \eta) \in T_{\text{epi}f}^\circ(x, f(x))$.

(2) 若 f 在 x 处是正则的, 作函数

$$g: \mathcal{B} \times R \rightarrow R, \quad g(x', \eta) = f(x') - \eta.$$

显然, g 在 $(x, f(x))$ 处是正则的, 并且 $\text{epi}f = \{(x, \eta) | g(x, \eta) \leq 0\}$. 由于 $0 \notin \partial^\circ g(x, f(x)) = \partial^\circ f(x) \times \{-1\}$, 由定理 4.4.9 可知集合 $\text{epi}f$ 在 $(x, f(x))$ 处也是正则的.

反之, 设

$$f'_+(x; d) = \liminf_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t},$$

我们证明 $Q_{\text{epi}f}(x, f(x)) = \text{epi}f'_+(x; \cdot)$ 成立. 事实上, 设 $(d, \eta) \in Q_{\text{epi}f}(x, f(x))$, 则对任给的 $\varepsilon > 0$, 存在 $t \in (0, \varepsilon)$, $p \in d + \varepsilon B$ 和 $\eta' \in \eta + (-\varepsilon, \varepsilon)$ 使得 $(x, f(x)) + t(p, \eta') \in \text{epi}f$, 即 $f(x) + t\eta' \geq f(x + tp)$, 从而有

$$\frac{f(x + tp) - f(x)}{t} \leq \eta' \leq \eta + \varepsilon.$$

令 $\varepsilon \rightarrow 0$, $t \rightarrow 0^+$, $p \rightarrow d$, 对上式取极限, 则得 $f'_+(x; d) \leq \eta$. 又对任

给的 $\delta > 0$, 由 $f'_+(x; d)$ 的定义, 则必存在 $t_k \rightarrow 0^+$, 使得

$$\frac{f(x + t_k d) - f(x)}{t_k} \leq f'_+(x; d) + \delta, \quad k = 1, 2, \dots$$

取 $s = f'_+(x; d) + \delta$, 有

$$f(x) + t_k s \geq f(x + t_k d), \quad k = 1, 2, \dots$$

由此, 据定义 4.4.2 知 $(d, f'_+(x; d) + \delta) \in Q_{\text{epi}f}(x, f(x))$. 由于 $\text{epi}f$ 在 $(x, f(x))$ 处是正则的, 由 (1) 的证明我们有

$$\text{epi}f'_+(x; \cdot) = \text{epi}f^0(x; \cdot).$$

于是, 对任意的 $d \in \mathcal{B}$ 有 $f'_+(x; d) = f^0(x; d)$, 从而 $f'_+(x; d) = f^0(x; d)$. 据此, 由定义 4.2.4 得 f 在 x 处是正则的. \square

推论 4.4.15 设 $f: \mathcal{B} \rightarrow R$ 是实值函数, $x \in \mathcal{B}$, 则 $x^* \in \partial^0 f(x)$ 当且仅当 $(x^*, -1) \in N_{\text{epi}f}^0(x, f(x))$.

证明 因为 $x^* \in \partial^0 f(x)$ 当且仅当对任意的 $d \in \mathcal{B}$ 有 $f^0(x; d) \geq \langle x^*, d \rangle$. 这意味着对任意的 $d \in \mathcal{B}$ 和任何 $\eta \geq f^0(x; d)$, 有

$$\langle (x^*, -1), (d, \eta) \rangle \leq 0.$$

根据定理 4.4.5, 它等价于 $(d, \eta) \in T_{\text{epi}f}^0(x, f(x))$, 由定义 4.4.1 即得 $(x^*, -1) \in N_{\text{epi}f}^0(x, f(x))$. \square

推论 4.4.15 告诉我们, 可以利用函数上图象的 G -法锥来定义它的 G -次梯度, 并且该函数可以不是 Lipschitz 的. 我们将在第 5 章进一步研究由切锥引进次梯度的问题.

利用 G -法锥, 可以给出一般函数的 G -次梯度的定义如下.

定义 4.4.4 设函数 $f: \mathcal{B} \rightarrow R \cup \{\pm \infty\}$ 在点 $x^0 \in \mathcal{B}$ 处是有限的. 若存在 $x^* \in \mathcal{B}^*$ 使得 $(x^*, -1) \in N_{\text{epi}f}^0(x^0, f(x^0))$, 则称 x^* 是 f 在点 x^0 处的 G -次梯度. f 在点 x^0 处的所有 G -次梯度的集合称为 f 在点 x^0 处的 G -次微分, 记作 $\partial^0 f(x^0)$. 若集合 $\text{epi}f$ 在点 $(x^0, f(x^0))$ 处是正则的, 则称 f 在点 x^0 处是正则的.

定理 4.4.16 设函数 $f: \mathcal{B} \rightarrow R \cup \{\pm \infty\}$ 在 $\bar{x} \in \mathcal{B}$ 处有限.

若对 \bar{x} 的某一邻域内任何一点 x 有 $f(\bar{x}) \leq f(x)$, 则 $\theta \in \partial^0 f(\bar{x})$.

证明 对任意的 $(d, \eta) \in T_{\text{epi}f}^0(\bar{x}, f(\bar{x}))$, 存在序列 $\{(d^k, \eta_k)\}$, $(d^k, \eta_k) \rightarrow (d, \eta)$ 和 $\{t_k\}$, $t_k \rightarrow 0^+$ ($k \rightarrow \infty$), 使得 $(\bar{x}, f(\bar{x})) + t_k(d^k, \eta_k) \in \text{epi}f$, 即有

$$f(\bar{x}) + t_k \eta_k \geq f(\bar{x}, t_k d^k).$$

由此, 对充分大的 k , 我们有 $f(\bar{x}) + t_k \eta_k \geq f(\bar{x}, t_k d^k) \geq f(\bar{x})$, 从而 $t_k \eta_k \geq 0$. 由 $t_k > 0$, 则有 $\lim_{k \rightarrow \infty} \eta_k = \eta \geq 0$, 于是对任意的 $(d, \eta) \in T_{\text{epi}f}^0(\bar{x}, f(\bar{x}))$, 有 $\langle (\theta, -1), (d, \eta) \rangle \leq 0$, 故 $\theta \in \partial^0 f(\bar{x})$. \square

定理 4.4.17 设集合 $S \subset \mathcal{B}$ 非空. 对于 S 上的指示函数

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathcal{B} \setminus S, \end{cases}$$

若 $x \in S$, 则 $\partial^0 \delta_S(x) = N_S^0(x)$. 并且, 函数 δ_S 在 x 处是正则的当且仅当 S 在 x 处是正则的.

证明 $x^* \in \partial^0 \delta_S(x)$ 等价于 $(x^*, -1) \in N_{\text{epi} \delta_S}^0(x, \theta)$. 显然有 $\text{epi} \delta_S = S \times [0, +\infty)$. 根据推论 4.4.8, 它等价于 $x^* \in N_S^0(x)$, $-1 \in N_{R_+}^0(\theta)$. \square

§ 4.5 有限维情况

本节讨论当 \mathcal{B} 为 Euclid 空间 R^n 时, G -次梯度和 G -次微分所特有的性质. 这时有 $\mathcal{B}^* = R^n$, 以及 $\partial^0 f(x) \subset R^n$.

一个重要的结果是, Lipschitz 函数在 R^n 的开子集上是几乎处处可微的. 由此, 函数在某点处的 G -次梯度可以由其附近的可微点逼近. 又几乎处处可微的点在 G -次可微点集上是稠密的. 以下记 Ω_f 表示 f 的不是 G -次可微的点组成的集合.

定理 4.5.1 设函数 $f: R^n \rightarrow R$ 在 $x \in R^n$ 处是 Lipschitz 的, $S \subset R^n$ 是 Lebesgue 测度为 0 的集合, 则有

$$\partial^0 f(x) = \text{co} \left\{ \lim_{k \rightarrow \infty} \nabla f(x^k) \mid x^k \rightarrow x, x^k \notin S, x^k \notin \Omega_f \right\}.$$

证明 设点列 $\{x^k\} \subset R^n$, $x^k \rightarrow x$ ($k \rightarrow \infty$), $x^k \notin S \cup \Omega_f$, $S \cup \Omega_f$ 的测度为 0, f 在 x^k 处 G -次可微. 由定理 4.1.1 可知 $\partial^0 f$ 在 x 附近是局部有界的, 又由定理 4.2.5, 则有 $\nabla f(x^k) \in \partial^0 f(x^k)$ ($k = 1, 2, \dots$). 根据 Bolzano-Weierstrass 定理, 序列 $\{\nabla f(x^k)\}$ 存在收敛的子列, 因为 $\partial^0 f(x)$ 是闭的, 所以 $\{\nabla f(x^k)\}$ 的聚点均属于 $\partial^0 f(x)$, 即

$$\left\{ \lim_{k \rightarrow \infty} \nabla f(x^k) \mid x^k \rightarrow x, x^k \notin S \cup \Omega_f \right\} \subset \partial^0 f(x).$$

由于 $\partial^0 f(x)$ 是凸集, 因此上式左边的凸包也包含于 $\partial^0 f(x)$. 下面证明对任意的 $\varepsilon > 0$ 和任意的 $d (\in R^n) \neq 0$, 有

$$f^0(x; d) - \varepsilon \leq \limsup_{y \rightarrow x} \{ \nabla f(y)^T d \mid y \rightarrow x, y \notin S \cup \Omega_f \}.$$

为此, 设 $a = \limsup_{y \rightarrow x} \{ \nabla f(y)^T d \mid y \rightarrow x, y \notin S \cup \Omega_f \}$. 由 a 的定义则存在 $\delta > 0$, 使得

$$y \in x + \delta B, y \notin S \cup \Omega_f, \nabla f(y)^T d \leq a + \varepsilon. \quad (4.5.1)$$

可选 δ 充分小使得 $\mu((S \cup \Omega_f) \cap (x + \delta B)) = 0$ ($\mu(\cdot)$ 是 R^n 中的测度函数), 并考虑线段

$$L_y = \left\{ y + td \mid 0 < t < \frac{\delta}{2 \|d\|} \right\}.$$

根据 Fubini 定理^[32], 对于任意的 $y \in x + (\delta/2)B$, $y \notin S \cup \Omega_f$ 和 $L_y \subset x + (\delta/2)B + td \subset x + \delta B$, 有 $\mu((S \cup \Omega_f) \cap L_y) = 0$.

设 $t \in \left(0, \frac{\delta}{2 \|d\|} \right)$, 则有

$$f(y + td) - f(y) = \int_0^t \nabla f(y + sd)^T d(ds).$$

因为 $\|y + sd - x\| \leq \delta$, $0 < s < t$, 由 (4.5.1) 和上式可得

$$f(y + td) - f(y) \leq t(a + \varepsilon)$$

$$\forall y \in x + (\delta/2)B, y \notin S \cup \Omega_f,$$

再由 f 是连续的, 可知上式对任意的 $y \in x + (\delta/2)B$, $t \in \left(0, \frac{\delta}{2\|d\|}\right)$ 均成立, 故有 $f^\circ(x; d) \leq a + \varepsilon$. 由此, 推得定理的结论成立. \square

显然有以下推论.

推论 4.5.2 设函数 $f: R^n \rightarrow R$ 在 $x \in R^n$ 处是 Lipschitz 的, 则

$$f^\circ(x; d) = \limsup_{y \rightarrow x} \{\nabla f(y)^T d \mid y \notin S \cup \Omega_f\}.$$

例 4.5.1 考虑函数 $f: R^2 \rightarrow R$, $f(x, y) = \max\{\min\{x, -y\}, y - x\}$. 设

$$S_1 = \{(x, y) \mid y \leq 2x, y \leq -x\},$$

$$S_2 = \{(x, y) \mid y \leq 2x, y \geq -x\}$$

和

$$S_3 = \{(x, y) \mid y \geq 2x, \text{或 } y \geq x/2\},$$

则 $S_1 \cup S_2 \cup S_3 = R^2$, $f(x, y)$ 等价于

$$f(x, y) = \begin{cases} x, & (x, y) \in S_1; \\ -y, & (x, y) \in S_2; \\ y - x, & (x, y) \in S_3. \end{cases}$$

设 S 是 S_1 、 S_2 和 S_3 的边界构成的集合, S 的测度为 0. 对于 $(x, y) \notin S$ 的点, f 是可微的, 并且 $\nabla f(x, y)$ 是 $(1, 0)$ 或 $(0, -1)$ 或 $(-1, 1)$, 故 $\partial^\circ f(0, 0)$ 是这三个点的凸包 (图 4.5.1). 由 $f(x, 0) = \max\{0, -x\}$, 则 $\partial_1^\circ f(0, 0) = [-1, 0]$, 又 $f(0, y) = \max\{0, y\}$, 故 $\partial_2^\circ f(0, 0) = [0, 1]$. 于是有

$$\partial^\circ f(0, 0) \neq \partial_1^\circ f(0, 0) \times \partial_2^\circ f(0, 0).$$

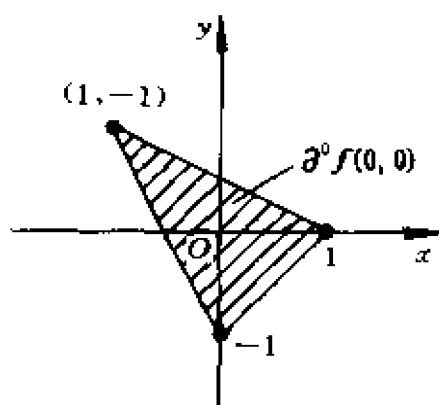


图 4.5.1

定理 4.5.3 设 $a \in R^n$, $b \in R^m$, 函数 $f: R^n \times R^m \rightarrow R$ 在点 $x = (a, b)$ 的凸邻域 $U \times V \subset R^n \times R^m$ 上是 Lipschitz 的, 对于每个 $a' \in U$, 函数 $f(a', \cdot)$ 在 V 上是凸的. 若 $(z, w) \in \partial^0 f(x)$, 则 $w \in \partial_2^0 f(a, b)$.

证明 由定理 4.5.1, 对于 $(z, w) \in \partial^0 f(x)$, 存在 $x^k = (a^k, b^k) \rightarrow x$ ($k \rightarrow \infty$), $x^k \notin \Omega_f$, 使得 $\nabla f(x^k) \rightarrow (z, w)$. 设 $\nabla f(x^k) = (z^k, w^k)$, 根据定理 4.2.6 和定理 4.2.11 知 $w^k \in \partial_2^0 f(a^k, b^k)$. 因为 $f(a^k, \cdot)$ 关于 b^k 是凸函数, 故对任意的 $d \in R^m$ 和充分大的 k , 有

$$f(a^k, b^k + d) - f(a^k, b^k) \geq \langle d, w^k \rangle.$$

对上式令 $k \rightarrow \infty$, 我们有

$$f(a, b + d) - f(a, b) \geq \langle d, w \rangle.$$

由此, 即得 $w \in \partial_2^0 f(a, b)$. \square

考虑 R^n 中集合 S 的距离函数:

$$d_S(x) = \inf_{y \in S} \|x - y\|, \quad x \in R^n,$$

其中 $\|x - y\| = \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}$, $x = (x_1, \dots, x_n)^T$, $y = (y_1, \dots, y_n)^T$.

定理 4.5.4 设集合 $S \subset R^n$ 非空. 若 $\nabla d_S(x)$ 存在且不为 θ , 则 $x \notin \text{cl}S$, 并且存在唯一的点 $x^0 \in \text{cl}S$, 使得

$$\nabla d_S(x) = \frac{x - x^0}{\|x - x^0\|}.$$

证明 反证: 假设 $x \in \text{cl}S$, 则对任意的 $d \in R^N$ 有

$$\begin{aligned} \nabla d_S(x)^T d &= \lim_{t \rightarrow 0^+} \frac{d_S(x + td) - d_S(x)}{t} \\ &= \lim_{t \rightarrow 0^+} \frac{d_S(x + td)}{t} \geqslant 0. \end{aligned}$$

由此得 $\nabla d_S(x) = \theta$, 它与已知相矛盾. 从 $x \notin \text{cl}S$, 则存在一点 $x^0 \in \text{cl}S$ 使得 $d_S(x) = \|x - x^0\|$. 设 $t \in (0, 1)$, 则显然有

$$d_S(x + t(x^0 - x)) = (1 - t) \|x - x^0\|,$$

于是

$$\frac{d_S(x + t(x^0 - x)) - d_S(x)}{t} = - \|x - x^0\|.$$

令 $t \rightarrow 0^+$, 得到

$$d'_S(x; x^0 - x) = \nabla d_S(x)^T (x^0 - x) = - \|x - x^0\|.$$

因为 d_S 是 1 秩 Lipschitz 的, 所以 $\|\nabla d_S(x)\| \leqslant 1$, 从而有

$$\nabla d_S(x) = \frac{x - x^0}{\|x - x^0\|}. \quad \square$$

定义 4.5.1 设集合 $S \subset \mathcal{B}$ 非空, $d \in \mathcal{B}$, $d \neq \theta$. 若存在唯一的点 $x^0 \in \text{cl}S$, 使得 $d = x - x^0$, 其中 $x \in \mathcal{B}$ 满足 $d_S(x) = \|x - x^0\|$, 则称 d 在点 x^0 处垂直于 S , 记作 $d \perp S$ (在 x^0 处).

定理 4.5.5 设集合 $S \subset R^n$ 非空, $d \in R^n$, $d \neq \theta$. 若 d 在 $x \in \text{cl}S$ 处垂直于 S , 则对任意的 $y \in \text{cl}S$, 有

$$\langle d, y - x \rangle \leqslant \frac{1}{2} \|y - x\|^2.$$

证明 由定义 4.5.1, 对任意的 $y \in \text{cl}S$, 有 $d = x' - x$ 和

$$\|x' - y\| \geq \|x' - x\| \quad \forall x' \in S.$$

从上式有

$$\langle x' - y, x' - y \rangle \geq \langle x' - x, x' - x \rangle = \|d\|^2.$$

据此, 得

$$\begin{aligned} & \|d\|^2 + 2\langle d, x - y \rangle + \|x - y\|^2 \\ &= \langle d + x - y, d + x - y \rangle \\ &= \langle x' - y, x' - y \rangle \geq \|d\|^2, \end{aligned}$$

所以结论成立. \square

定理 4.5.6 设集合 $S \subset R^n$ 非空, $x \in \text{cl}S$, 则

$$\begin{aligned} \partial^0 d_S(x) = \text{co} \Big(\{ \mathbf{0} \} \cup \Big\{ d = \lim_{k \rightarrow \infty} \frac{d^k}{\|d^k\|} \Big| d^k \perp \\ S(\text{在 } x^k \text{ 处}), x^k \rightarrow x, d^k \rightarrow \mathbf{0} \Big\} \Big). \end{aligned}$$

证明 设

$$\begin{aligned} A = \text{co} \Big(\{ \mathbf{0} \} \cup \Big\{ d = \lim_{k \rightarrow \infty} \frac{d^k}{\|d^k\|} \Big| d^k \perp \\ S(\text{在 } x^k \text{ 处}), x^k \rightarrow x, d^k \rightarrow \mathbf{0} \Big\} \Big), \end{aligned}$$

由定理 4.5.4 知 $\nabla d_S(x)$ 或为 $\mathbf{0}$ 或为单位向量. 因此, 根据定理 4.5.1 得知 $\partial^0 d_S(x) \subset A$.

显然, d_S 在 x 处达到极小值, 故 $\mathbf{0} \in \partial^0 d_S(x)$. 下设 d^k 在 x^k 处垂直于 S , 即 $d^k = y^k - x^k$, $d_S(y^k) = \|y^k - x^k\|$, 并且 $x^k \rightarrow x$, $d^k \rightarrow \mathbf{0}$. 利用定理 4.3.5, 我们有

$$d_S(y^k) - d_S(x^k) \in \langle \partial^0 d_S(\xi^k), y^k - x^k \rangle,$$

其中 $\xi^k \in (x^k, y^k)$. 因为 $d_S(x^k) = 0$, 上式可写成

$$1 \in \left\langle \partial^0 d_S(\xi^k), \frac{y^k - x^k}{\|y^k - x^k\|} \right\rangle.$$

又因为 $\|\partial^0 d_S(x)\| \leq 1$, 从而可得 $\frac{d^k}{\|d^k\|} \in \partial^0 d_S(\xi^k)$. 由 $x^k \rightarrow x$, $d^k \rightarrow 0$, 有 $\xi^k \rightarrow x$, 依据定理 4.2.1 的(2)得 $d \in \partial^0 d_S(x)$, 因此 $A \subset \partial^0 d_S(x)$. \square

推论 4.5.7 设集合 $S \subset R^n$ 非空. 若 x 属于 $\text{cl}S$ 的边界点, 则 $\partial^0 d_S(x)$ 包含非零点.

证明 因 x 属于 $\text{cl}S$ 的边界点, 则对任意的 $\varepsilon > 0$, 存在点 $y_\varepsilon \in R^n$, 使得 $\|y_\varepsilon - x\| \leq \varepsilon$, 并且 $y_\varepsilon \notin \text{cl}S$. 由定理 4.5.4 知, 存在垂直向量 d_ε 使得 $\nabla d_S(y_\varepsilon) = \frac{d_\varepsilon}{\|d_\varepsilon\|}$. 令 $\varepsilon \rightarrow 0^+$, 有 $y_\varepsilon \rightarrow x$, 则 $\frac{d_\varepsilon}{\|d_\varepsilon\|} \rightarrow d \in \partial^0 d_S(x)$, $\|d\| = 1$. \square

推论 4.5.8 设集合 $S \subset R^n$ 非空. 若 x 属于 $\text{cl}S$ 的边界点, 则 $N_S^0(x)$ 包含非零点.

证明 由定理 4.4.2 和推论 4.5.7 即得证. \square

定理 4.5.9 设集合 $S \subset R^n$ 非空, $x \in \text{cl}S$, 并且

$$A = \text{co} \left(\{0\} \cup \left\{ d = \lim_{k \rightarrow \infty} \frac{d^k}{\|d^k\|} \mid d^k \perp S(\text{在 } x^k \text{ 处}), x^k \rightarrow x, d^k \rightarrow 0 \right\} \right).$$

则 $N_S^0(x) = \bigcup_{\lambda \geq 0} \lambda A$.

证明 根据定理 4.4.2, $N_S^0(x)$ 是闭凸锥, 再由定理 4.5.6 即可得证. \square

定理 4.5.10 设 $S \subset R^n$ 是闭集, $x \in S$, 则 $d \in \text{int}T_S^0(x)$ 当且仅当存在 $\varepsilon > 0$ 使得

$$d_S(y + tp) \leq d_S(y) \quad \forall y \in x + \varepsilon B,$$

$$p \in d + \varepsilon B, t \in [0, \varepsilon). \quad (4.5.2)$$

证明 若存在 d 满足 (4.5.2), 则对任意的 $p \in d + \varepsilon B, x \in S$, 可以找到点列 $\{x^k\} \subset (x + \varepsilon B) \cap S, x^k \rightarrow x, \{t_k\} \in (0, \varepsilon), t_k \rightarrow 0^+ (k \rightarrow \infty)$, 对充分大的 k 有

$$d_S(x^k + t_k p) \leq d_S(x^k) = 0.$$

因此, 对充分大的 k 我们有 $x^k + t_k p \in S$. 再由定理 4.4.7 知 $p \in T_S^0(x)$, 故得 $d \in \text{int} T_S^0(x)$.

设 $d \in \text{int} T_S^0(x)$, 则对任一非零的 $z \in N_S^0(x)$ 有 $d^T z < 0$ (否则, 若 $d^T z = 0$, 则在 d 的任一邻域中存在一点 \tilde{p} 使得 $\tilde{p}^T z > 0$, 但另一方面, 从 $p \in T_S^0(x)$ 有 $p^T z \leq 0$, 导致矛盾). 因此, 存在 $\gamma > 0$ 使得

$$p^T z \leq -\gamma \|z\| \quad \forall z \in N_S^0(x).$$

利用定理 4.4.2, 从上式可推知

$$p^T z \leq -\gamma \|z\| \quad \forall z \in \partial^0 d_S(x). \quad (4.5.3)$$

现在证明: 存在某个 $\bar{w} > 0$, 使得对任意的 $p \in d + \bar{w}B$ 有

$$p^T \nabla d_S(y) \leq 0 \quad \forall y \in x + \bar{w}B. \quad (4.5.4)$$

反之, 假若 (4.5.4) 不成立, 则存在点列 $\{y^k\} \subset R^n, y^k \rightarrow x (k \rightarrow \infty)$, 使得

$$p^T \nabla d_S(y^k) > 0.$$

由定理 4.5.4, $\nabla d_S(y^k)$ 是非零的单位向量, 当 $k \rightarrow \infty$ 时, 它收敛到点 $z \in \partial^0 d_S(x)$. 对上式取极限得

$$p^T z \geq 0,$$

导致与 (4.5.3) 矛盾.

当 $y \in x + \varepsilon B, p \in d + \varepsilon B, t \in [0, \varepsilon)$ 时, 选取 ε 使得 $y + tp \in x + wB$, 注意, 函数 d_S 是几乎处处可微的, 固定 p , 有

$$d_s(y + tp) - d_s(y) = \int_0^1 \nabla d_s(y + sp)^T p ds.$$

最后, 利用(4.5.4), 对几乎所有的 $y \in x + \varepsilon B$, 有 $d_s(y + tp) \leq d_s(y)$. 由 d_s 的连续性知(4.5.2)成立. \square

推论 4.5.11 设 $S \subset R^n$ 是闭集, $x \in S$, 则 $T_s^H(x) = \text{int} T_s^0(x)$.

证明 根据定理 4.5.10 知, 由于 $d \in T_s^H(x)$, 从(4.5.2)有 $d \in \text{int} T_s^0(x)$. 反之, 设 $d \in \text{int} T_s^0(x)$, 由定理 4.4.12 得 $d \in T_s^H(x)$. \square

推论 4.5.12 设 $S \subset R^n$ 是闭集, $x \in S$. 若 $\text{int} T_s^0(x) \neq \emptyset$, 则对任意的 $x^k \rightarrow x$, $\xi^k \in N_s(x^k)$, $\xi^k \rightarrow \xi$ ($k \rightarrow \infty$), 有 $\xi \in N_s^0(x)$.

证明 由推论 4.5.11 容易推得. \square

§ 4.6 Lipschitz 向量函数的 G -次微分

最后, 考虑满足 Lipschitz 条件的向量函数的 G -次微分问题.

定义 4.6.1 设集合 $S \subset R^n$ 非空, $f: X \rightarrow R^m$ 是向量函数 ($m \geq 1$), $f(x) = (f_1(x), \dots, f_m(x))^T$. 若每一个 f_i ($i = 1, \dots, m$) 在点 $x^0 \in S$ 处是 $(\gamma$ 秩) Lipschitz 的, 则称向量函数 f 在点 x^0 处是 Lipschitz 的. 若每一个 f_i ($i = 1, \dots, m$) 是 S 上的 $(\gamma$ 秩) Lipschitz 函数, 则称 f 是 S 上的 $(\gamma$ 秩) Lipschitz 向量函数.

设 $f(x) = (f_1(x), \dots, f_m(x))^T$ 在点 $x \in R^n$ 处是 Lipschitz 的, 这时每一个 f_i ($i = 1, \dots, m$) 是几乎处处可微的. 相应地, 我们说 f 在点 x 的邻域内是几乎处处可微的. 用 Ω_f 表示 f 的不可微点组成的集合. 若各 f_i ($i = 1, \dots, m$) 在 x 处的偏导数均存在, 则用 $Jf(x)$ 表示由这些偏导数构成的 $m \times n$ 阶 Jacobi 矩阵.

定义 4.6.2 设 $x^0 \in R^n$, $x^k \in R^n$ ($k = 1, 2, \dots$), $f: R^n \rightarrow R^m$ 是 Lipschitz 向量函数. 记

$$d^0 f(x^0) = \text{co} \left\{ \lim_{k \rightarrow \infty} Jf(x^k) \mid x^k \rightarrow x^0, x^k \notin \Omega_f \right\},$$

则称 $x^* \in \partial^0 f(x^0)$ 是 f 在点 x^0 处的 G -次梯度, 称 $\partial^0 f(x^0)$ 是 f 在点 x^0 处的 G -次微分. 若 $\partial^0 f(x^0) \neq \emptyset$, 则称 f 在点 x^0 处是 G -次可微的.

对于 $m \times n$ 阶矩阵 $M \in R^{m \times n}$, 定义它的范数:

$$\|M\|_{m \times n} = \left\{ \sum_{i=1}^m \|r_i\|^2 \mid r_i \text{ 是 } M \text{ 的第 } i \text{ 行向量} \right\}^{1/2},$$

并记 $B_{m \times n} \in R^{m \times n}$ 表示开单位球.

定理 4.6.1 设 $f: R^n \rightarrow R^m$ 在 $x \in R^n$ 处是 Lipschitz 的, $f(x) = (f_1(x), \dots, f_m(x))^T$.

(1) $\partial^0 f(x) \subset R^{m \times n}$ 是非空紧凸集.

(2) $\partial^0 f(x)$ 在 x 处是闭的, 即若 $x^k \in R^n (k = 1, 2, \dots)$, $x^k \rightarrow x$, $z^k \in \partial^0 f(x^k)$, $z^k \rightarrow z (k \rightarrow \infty)$, 则 $z \in \partial^0 f(x)$.

(3) 对任意的 $\varepsilon > 0$, 存在 $\delta > 0$ 使得

$$\partial^0 f(y) \subset \partial^0 f(x) + \varepsilon B_{m \times n} \quad \forall y \in x + \delta B.$$

(4) 若 $f_i (i = 1, \dots, m)$ 在 x 处是 γ_i 秩 Lipschitz 的, 则 f 在 x 处是 $\gamma = \|\gamma_1, \dots, \gamma_m\|^T$ 秩 Lipschitz 的, 并且 $\partial^0 f(x) \subset \gamma \text{cl} B_{m \times n}$.

(5) $\partial^0 f(x) \subset \partial^0 f_1(x) \times \partial^0 f_2(x) \times \dots \times \partial^0 f_m(x)$, 表达式右端意即矩阵的第 i 行的元素属于 $\partial^0 f_i(x)$. 当 $m = 1$ 时, $\partial^0 f(x) = \partial^0 f_1(x)$.

证明 (1)~(5)都可由定理 4.5.1 推得. \square

设 $S \subset R^n$ 是测度为 0 的集合. 在定义 4.6.1 中选取点列 $\{x^k\} \not\subset S$, 可得到另一个广义 Jacobi 矩阵:

$$\partial_s^0 f(x) = \text{co} \left\{ \lim_{k \rightarrow \infty} Jf(x^k) \mid x^k \rightarrow x, x^k \notin S \cap \Omega_f \right\}.$$

事实上, 可以证明 $\partial_s^0 f(x)$ 与 $\partial^0 f(x)$ 是一致的.

定理 4.6.2 设 $f: R^n \rightarrow R^m$ 在 $x \in R^n$ 处是 Lipschitz 的, 则对任意的 $d \in R^n$ 和任意的 $p \in R^m$, 有

$$\partial^0 f(x)d = \partial_S^0 f(x)d,$$

$$\partial^0 f(x)^T p = \partial_S^0 f(x)^T p.$$

证明 由定义显然有 $\partial_S^0 f(x) \subset \partial^0 f(x)$. 对任意的 $u \in R^n$, 作 $\partial^0 f(x)d$ 的支撑函数 σ_1 和 $\partial_S^0 f(x)d$ 的支撑函数 σ_2 :

$$\begin{aligned}\sigma_1(u) &= \limsup \{ \langle u, Jf(y)d \rangle \mid y \rightarrow x, y \notin \Omega_f \} \\ &= \limsup \{ \langle Jf(y)^T u, d \rangle \mid y \rightarrow x, y \notin \Omega_f \},\end{aligned}$$

$$\begin{aligned}\sigma_2(u) &= \limsup \{ \langle u, Jf(y)d \rangle \mid y \rightarrow x, y \notin S \cap \Omega_f \} \\ &= \limsup \{ \langle Jf(y)^T u, d \rangle \mid y \rightarrow x, y \notin S \cap \Omega_f \}.\end{aligned}$$

再作函数 $g: R^n \rightarrow R$, $g(y) = \langle u, f(y) \rangle$, 注意 ∇g 存在, 并当 $y \notin \Omega_f$ 时有 $\nabla g(y)^T = Jf(y)^T u$. 由此, 根据定理 4.5.1 得

$$\begin{aligned}\sigma_1(u) &= \limsup \{ \nabla g(y)^T d \mid y \rightarrow x, y \notin \Omega_f \} = g^0(x; d) \\ &= \limsup \{ \nabla g(y)^T d \mid y \rightarrow x, y \notin S \cap \Omega_f \} = \sigma_2(u).\end{aligned}$$

由定理 4.2.11, 即知定理的结论成立. \square

以下是中值定理和链式法则.

定理 4.6.3 设 $f: R^n \rightarrow R^m$ 是开凸集 $S \in R^n$ 上的 Lipschitz 向量函数, $x, y \in S$, 则

$$f(y) - f(x) \in \text{co} \partial^0 f([x, y]) (y - x).$$

证明 只要证明存在 $z \in \partial^0 f(\xi)$, $\xi \in [x, y]$, 使得 $f(y) - f(x) \in \text{co}(z(y - x))$. 当固定 $x, y \in \Omega_f$, 线段 $[x, y]$ 与 Ω_f 的交集的测度为 0, 我们有

$$f(y) - f(x) = \int_0^1 Jf(x + t(y - x)) (y - x) dt,$$

因此, $f(y) - f(x) \in \text{co} \partial^0 f([x, y]) (y - x)$. \square

定理 4.6.4 设 $h: R^n \rightarrow R^m$ 在 $x \in R^n$ 处是 Lipschitz 的, $g: R^m \rightarrow R$ 在 $h(x)$ 处是 Lipschitz 的, $f = g \circ h$, 则 f 在 x 处是

Lipschitz 的, 并且有

$$\partial^0 f(x) \subset \text{co} \{ \partial^0 g(h(x)) \circ \partial^0 h(x) \}. \quad (4.6.1)$$

再若 g 在 x 处是严格可微的, 则 (4.6.1) 成为等式.

证明 容易证明 f 在 x 附近是 Lipschitz 的. 设对任意的 $\epsilon > 0$, $y \in x + \epsilon B$, 有

$$f^0(x; d) \leq \frac{f(y + td) - f(y)}{t} + \epsilon,$$

根据定理 4.3.5, 我们有

$$f(y + td) - f(y) = \zeta^T (h(y + td) - h(y)),$$

其中 $\zeta \in \partial^0 g(\xi)$, $\xi \in [h(y), h(y + td)]$. 再由定理 4.6.2 可知有 $p \in \text{co} \partial^0 h[y, y + td]d$,

$$h(y + td) - h(y) = tp.$$

由此, 存在 $z \in \partial^0 h[y, y + td]$ 使得 $\zeta^T p = \zeta^T (zd)$, 从而

$$f^0(x; d) \leq \zeta^T (zd) + \epsilon, \quad \zeta^T \in \partial^0 g(\xi), \quad \xi \in [h(y), h(y + td)],$$

$$z \in \partial h[y, y + td], \quad y \in x + \epsilon B. \quad (4.6.2)$$

可以选择点列 $\{y^k\} \subset R^n$, $y^k \rightarrow x$, $\{t_k\} \in (0, \infty)$, $t_k \rightarrow 0$ ($k \rightarrow \infty$) 使得 (4.6.2) 成立. 显然, 对应地有 $\xi^k \rightarrow h(x)$, $\zeta_k \rightarrow \zeta_0 \in \partial^0 g(h(x))$, $z_k \rightarrow z_0 \in \partial^0 h(x)$, 由 (4.6.2) 得到

$$f^0(x; d) \leq \zeta_0^T z_0 d + \epsilon.$$

由 ϵ 的任意性, 有 $f^0(x; d) \leq \zeta_0^T z_0 d$, 再由定理 4.2.11 得知 (4.6.1) 成立.

现设 g 在 $h(x)$ 处是严格可微的, $D_S g(h(x)) = \zeta$, 对任给的 $\epsilon > 0$, 使得 h 在 $x + \epsilon B$ 上是 γ 秩 Lipschitz 的. 选择 $\delta \in (0, \epsilon)$, 则对任意的 $y \in x + \delta B$, h 在 y 处是 Lipschitz 的, g 在 $h(y)$ 处是 Lipschitz 的. 由定理 4.6.1 的 (3) 选取充分小的 δ , 可使

$$\partial^0 g(h(y)) \subset \xi + \frac{\varepsilon}{\gamma} B. \quad (4.6.3)$$

注意 $\partial^0 g(h(x)) = \{\xi\}$, 对任何向量 $d \in R^n$, 当 $y \notin \Omega_f \cup \Omega_h$ 时, 有

$$\nabla f(y)^T d = \lim_{t \rightarrow 0^+} \frac{g(h(y + td)) - g(h(y))}{t}.$$

于是, 由定理 4.6.3 可知, 存在 $\xi' \in \partial^0 g(\xi')$ 和 $\xi' \in [h(y), h(y + td)]$, 使得

$$\nabla f(y)^T d = \lim_{t \rightarrow 0^+} (\xi')^T \frac{h(y + td) - h(y)}{t}.$$

取 $\{\xi'\}$ 的收敛的子列 $\xi' \rightarrow \xi$, 因为 h 在 y 处可微, $\partial^0 g(h(y)) \circ Jh(y)$ 是凸的, 从上式得

$$\nabla f(y)^T d \in \partial^0 g(h(y)) \circ Jh(y)d. \quad (4.6.4)$$

由 (4.6.3) 和 (4.6.4), 对任意的 $y \in x + \delta B$, $y \notin \Omega_f \cup \Omega_h$, 按定理 4.6.1 的 (1), 我们有

$$\begin{aligned} \nabla f(y) &\in \partial^0 g(h(y)) \circ Jh(y) \\ &\subset \left\{ \xi + \frac{\varepsilon}{\gamma} B \right\} \circ J(h) \subset \xi \circ Jh(y) + \varepsilon B. \end{aligned} \quad (4.6.5)$$

设对任意的 $d \in R^n$, 有

$$\begin{aligned} q &= \max \xi^T \partial^0 h(x) d \\ &= \limsup \{ \xi'^T Jh(y) d \mid y \rightarrow x, y \notin \Omega_h \} \\ &\geq \limsup \{ \xi'^T Jh(y) d \mid y \rightarrow x, y \notin \Omega_f \cup \Omega_h \}. \end{aligned}$$

从 (4.6.5) 和推论 4.5.2, 我们有

$$q \geq \limsup \{ \nabla f(y)^T d \mid y \rightarrow x, y \in \Omega_f \cup \Omega_h \},$$

即

$$q \geq f^0(x; d).$$

再由 (4.6.5) 得 $q \leq f^0(x; d)$, 故 (4.6.1) 成为等式. \square

推论 4.6.5 设 $h: R^n \rightarrow R^m$ 在 $x \in R^n$ 处是 Lipschitz 的, $g: R^m \rightarrow R$ 在 $h(x)$ 处是 Lipschitz 的, 则对任意的 $d \in R^n$ 有

$$\partial^0(g \circ h)(x)d \subset \text{co} \{ \partial^0 g(h(x)) \circ \partial^0 h(x)d \}. \quad (4.6.6)$$

再若 g 在 $h(x)$ 处是连续可微的, 则 (4.6.6) 成为等式.

证明 设对任意的 $z, y \in R^n$, $g_1(y) = \langle z, y \rangle$, 记 $f_1 = g \circ h$, $f = g_1 \circ f_1$, 应用定理 4.6.4 得

$$\begin{aligned} z^T \partial^0(f_1(x))d &= z^T \partial^0(g \circ h)(x)d \\ &= \partial^0(z^T [g \circ h])(x)d \\ &\subset \text{co} \{ \partial^0(z^T g)(h(x)) \circ \partial^0 h(x)d \}. \end{aligned}$$

再根据定理 4.6.4, 上式为

$$\begin{aligned} z^T \partial^0 g(h(x))d &\subset \text{co} \{ z^T \partial^0 g(h(x)) \circ \partial^0 h(x)d \} \\ &= z^T \text{co} \{ \partial^0 g(h(x)) \circ \partial^0 h(x)d \}. \end{aligned}$$

由 z 的任意性, 故 (4.6.6) 成立.

若 g 在 $h(x)$ 处是连续可微的, 则显然有

$$\partial^0(g \circ h)(x) \supset \partial^0 g(h(x)) \circ \partial^0 h(x).$$

因为 $\Omega_{g,h} \subset \Omega_h$, 再由定理 4.6.4 得知 (4.6.6) 成为等式. \square

第 5 章 函数的切导数 和切微分

在第 3 章和第 4 章, 我们知道切锥与次梯度有着紧密的联系. 因为通过与切锥对应的法锥, 可以相应地建立函数的次梯度和次微分. 我们还看到, 对于函数而言, 它的上图象的切锥与方向导数的上图象是一致的. 因此, 借助切锥也是引进方向导数的一个重要途径.

本章将再介绍几种切锥以及它们的有关性质. 由此, 引进一般函数的方向导数、次梯度和次微分概念, 并得到一些相应的结果.

§ 5.1 凸集的切锥

在 § 3.2 中的定义 3.2.2, 我们曾利用指示函数的次微分, 引进了线性拓扑空间中凸集在一点处的法锥和切锥的概念. 现在进而给出在 Banach 空间 \mathcal{B} 中, 凸集 $S \subset \mathcal{B}$ 在点 $x \in S$ 处的切锥 $T_S(x)$ 和法锥 $N_S(x)$ 的等价形式, 并且讨论它们的有关性质.

设 \mathcal{B}^* 是 \mathcal{B} 的对偶空间.

定理 5.1.1 设 $S \subset \mathcal{B}$ 是非空凸集, $x \in S$, 则

$$T_S(x) = \text{cl} \left(\bigcup_{\lambda > 0} \lambda(S - x) \right), \quad (5.1.1)$$

$$N_S(x) = \{x^* \in \mathcal{B}^* \mid \langle x^*, d \rangle \leq 0 \quad \forall d \in T_S(x)\}. \quad (5.1.2)$$

证明 记 (5.1.1) 和 (5.1.2) 右端的集合分别为 T 和 N . 先证

(5.1.1). 设 $x^* \in T^-$, 由定义 1.4.6 的(3), 则对任意的 $d \in T$ 有 $\langle x^*, d \rangle \leq 0$, 从而对任意的 $y \in S$ 有 $\langle x^*, y - x \rangle \leq 0$. 据此, 按定义 3.2.2 中的(1)知 $x^* \in N_S(x)$, 于是 $T^- \subset N_S(x)$. 由注 3.2.4、注 1.4.3 和注 1.4.4, 则得 $T_S(x) = N_S(x) \subset (T^-)^- = T$. 反之, 设 $d \in T$, 则存在 $\lambda_k > 0$, $y^k \in S (k = 1, 2, \dots)$, 使得 $d = \lim_{k \rightarrow \infty} \lambda_k (y^k - x)$. 任取 $x^* \in N_S(x)$, 由定义 3.2.2 的(1), 则对任意的 k , 有 $\langle x^*, y^k - x \rangle \leq 0$, 从而 $\langle x^*, \lambda_k (y^k - x) \rangle \leq 0$. 令 $k \rightarrow \infty$, 得 $\langle x^*, d \rangle \leq 0$. 于是, 按定义 1.4.6 的(3)知 $d \in N_S(x)$. 因此, 再由注 3.2.4 得到 $T \subset N_S(x)^- = T_S(x)$.

为证(5.1.2), 设 $x^* \in N$, 由(5.1.1)可知对任意的 $y \in S$, 有 $\langle x^*, y - x \rangle \leq 0$, 故按定义 3.2.2 的(1), 得 $x^* \in N_S(x)$. 反之, 设 $x^* \in N_S(x)$, 根据定义 3.2.2 的(1)有

$$\langle x^*, y - x \rangle \leq 0 \quad \forall y \in S.$$

由此, 设 $d = \lim_{k \rightarrow \infty} \lambda_k (y^k - x) \in T_S(x)$, 其中 $\lambda_k > 0$, $y^k \in S (k = 1, 2, \dots)$, 则对任意的 k , 有 $\langle x^*, y^k - x \rangle \leq 0$, 从而 $\langle x^*, d \rangle \leq 0$. 于是, 得到 $x^* \in N$. \square

注 5.1.1 (5.1.1)和(5.1.2)分别是 Banach 空间中凸集的切锥和法锥的等价形式. 设 $B \subset \mathcal{B}$ 是开单位球, 由(5.1.1)容易知道 $d \in T_S(x)$ 等价于: 对任意的 $\varepsilon > 0$, 存在 $p \in d + \varepsilon B$ 和 $t > 0$ 使得 $x + tp \in S$. 它也等价于: 存在序列 $\{d^k\} \subset S$, $d^k \rightarrow d$ 和 $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0 (k \rightarrow \infty)$, 使得对任意的 $k = 1, 2, \dots$, 有 $x + t_k d^k \in S$. 此外, 它也等价于下式成立:

$$\liminf_{t_k \rightarrow 0} \frac{1}{t_k} d_S(x + t_k d^k) = 0,$$

其中 $d^k \rightarrow d (k \rightarrow \infty)$.

定理 5.1.2 设 $S \subset \mathcal{B}$ 是非空凸集, $x \in S$.

(1) $T_S(x) = T_{\text{cl}S}(x)$.

(2) $T_S(x) = \mathcal{B} (\forall x \in \text{int}S)$.

(3) $S \subset x + T_S(x)$.

证明 由(5.1.1)直接可得证. \square

定理 5.1.3 设 $S \subset \mathscr{E}$ 是闭凸集, $x \in S$, 则 $N_S(x)$ 是闭集.

证明 设点列 $\{x_k\} \subset T_S(x)$, $x_k \rightarrow x$, $\{x_k^*\} \in N_S(x_k)$, $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$). 根据定义 3.2.2 的(2), 对任意的 $d \in S$, 有 $\langle x_k^*, d \rangle \leq \langle x_k^*, x_k \rangle$. 令 $k \rightarrow \infty$, 得

$$\langle x^*, d \rangle \leq \langle x^*, x \rangle \quad \forall d \in S,$$

故 $x^* \in N_S(x)$. \square

定理 5.1.4 设 $S \subset \mathscr{E}$ 是凸集. 若 $\text{int}S \neq \emptyset$, 则

$$\text{int}T_S(x) = \bigcup_{\lambda > 0} \lambda(\text{int}S - x) \quad \forall x \in S, \quad (5.1.3)$$

并且映射: $x \mapsto \text{int}T_S(x)$ 的图象是开集.

证明 显然 $\bigcup_{\lambda > 0} \lambda(\text{int}S - x)$ 是开集, 并且 $\bigcup_{\lambda > 0} \lambda(\text{int}S - x) \subset \text{int}T_S(x)$. 另外, 因为 $\text{int}T_S(x) = \text{int}C_S(x)$ (其中 $C_S(x) = \bigcup_{\lambda > 0} \lambda(S - x)$), 设 $d \in \text{int}C_S(x)$, 可取 $\delta > 0$ 使 $d + \delta B \subset C_S(x)$. 若 $d \notin S - x$. 设 $x^0 \in \text{int}S$, $d^0 = x^0 - x$, 有 $d - \frac{\delta d^0}{\|d^0\|} \in C_S(x)$, 则存在 $\lambda > 0$ 使得

$$x + \frac{1}{\lambda} \left(d - \frac{\delta d^0}{\|d^0\|} \right) \in S.$$

令 $\alpha = \frac{\delta}{\delta + \lambda \|d^0\|}$, 我们有

$$\lambda x + (1 - \alpha)d = \alpha \lambda x^0 + (1 - \alpha) \left(\lambda x + \left(d - \frac{\delta d^0}{\|d^0\|} \right) \right) \in S.$$

从上式可得 $d \in \frac{\lambda}{(1 - \alpha)}(\text{int}S - x)$, 因而(5.1.3)成立.

设 $d^0 \in \text{int}T_S(x_0)$, 则存在某个 $\lambda_0 > 0$, 使 $d^0 \in \lambda_0(\text{int}S - x^0)$, 因此存在 $\varepsilon > 0$ 使得

$$x^0 + \lambda_0 d^0 + \varepsilon B = x^0 + \lambda_0 \left(d^0 + \frac{\varepsilon}{\lambda_0} B \right) \subset \text{int}S.$$

取 $x \in x^0 + \frac{\varepsilon}{2}B$ 和 $d \in d^0 + \frac{\varepsilon}{2\lambda_0}B$, 由上式得

$$x + \lambda_0 d \in x^0 + \lambda_0 d^0 + \varepsilon B \subset \text{int} S,$$

从而 $d \in \text{int} T_S(x)$. 由此, $x \rightarrow \text{int} T_S(x)$ 的图象是开集. \square

定理 5.1.5 设 \mathcal{H} 是 Hilbert 空间, $B \subset \mathcal{H}$ 是闭单位球, $x \in B$.

(1) 若 $x \in \text{int} B$, 则 $T_B(x) = \mathcal{H}$.

(2) 若 $x \in \text{int} B$, 则 $N_B(x) = \{0\}$.

证明 由定理 5.1.1 可得证. \square

设 S 表示集合 $S \subset \mathcal{B}$ 的负对偶集(或称极化集), 即 $S^\circ = \{x^* \in \mathcal{B}^* \mid \langle x^*, x \rangle \leq 0 \quad \forall x \in S\}$ (若 $S = K$ 是凸锥, 由定义 1.4.6 的(3), K° 是 K 的负对偶锥). S^\perp 表示集合 S 的垂直集, 即

$$S^\perp = \{x^* \in \mathcal{B}^* \mid \langle x^*, x \rangle = 0 \quad \forall x \in S\}.$$

定理 5.1.6 设 $K \subset \mathcal{B}$ 是闭凸锥, $x \in \mathcal{B}$, 则 $N_K(x) = K^\circ \cap \{x\}^\perp$, 即 $d \in T_K(x)$ 当且仅当对任意的 $x^* \in K^\circ$, 满足 $\langle x^*, x \rangle = 0$ 使得 $\langle x^*, d \rangle \leq 0$. 再若 $S \subset \mathcal{B}$ 是闭线性子空间, 则 $T_S(x) = S$, $N_S(x) = S^\perp$.

证明 显然有 $K^\circ \cap \{x\}^\perp \subset N_K(x)$. 反之, 若 $x^* \in N_K(x)$, 根据定义 3.2.2 的(1)有 $\langle x^*, x \rangle = \max_{d \in K} \langle x^*, d \rangle$, 因为 K 是锥, 所以有 $\langle x^*, x \rangle = 0$ 和 $\langle x^*, d \rangle \leq 0 (\forall d \in K)$. \square

设 \mathcal{D} 是 Banach 空间, $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是线性映射, $\text{Ker} \varphi = \{x \in \mathcal{B} \mid \varphi(x) = 0\}$ 是 φ 的核.

定理 5.1.7 设 $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是线性映射, $S = \varphi^{-1}(y) \subset \mathcal{B}$ 是仿射子空间. 若 $\varphi(x) = y$, 则 $T_S(x) = \text{Ker} \varphi$.

证明 若 $d \in \text{Ker} \varphi$, 则 $d + x \in \varphi^{-1}(y) = S$, 因此 $d \in T_S(x)$. 反之, 设 $d^k = \lambda_k(x^k - x)$, $x^k \in S$, $d^k \rightarrow d \in T_S(x) (k \rightarrow \infty)$, 则 $\varphi(d^k) = \lambda_k \varphi(x^k - x) = 0$. 由此, 得 $d^k \in \text{Ker} \varphi$, 从而 $d \in \text{Ker} \varphi$. \square

定理 5.1.8 设集合

$$S = \left\{ \mathbf{x} \in R_+^n \mid \mathbf{x} = (x_1, \dots, x_n)^T, \sum_{i=1}^n x_i = 1 \right\},$$

$$I(\mathbf{x}) = \{i = 1, \dots, n \mid x_i = 0\},$$

$$\mathbf{d} = (d_1, \dots, d_n)^T.$$

(1) $\mathbf{d} \in T_{R_+^n}(\mathbf{x})$ 当且仅当对任何 $i \in I(\mathbf{x})$ 有 $d_i \geq 0$.

(2) $\mathbf{d} \in T_S(\mathbf{x})$ 当且仅当对任何 $i \in I(\mathbf{x})$ 有 $d_i \geq 0$, 并且

$$\sum_{i=1}^n d_i = 0.$$

证明 (1) 由定理 5.1.6 得, 当 $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$ 时, 有

$$\sum_{i=1}^n x_i^* x_i = \sum_{i \notin I(\mathbf{x})} x_i^* x_i = 0,$$

从而当 $i \notin I(\mathbf{x})$ 时有 $x_i^* = 0$. 据此, 即得 $\mathbf{d} \in T_{R_+^n}(\mathbf{x})$ 当且仅当

$$\sum_{i \notin I(\mathbf{x})} x_i^* d_i \leq 0 \quad (\forall x_i^* \in R^n), \text{ 即对任何 } i \in I(\mathbf{x}), \text{ 有 } d_i \geq 0.$$

(2) 对任何 $i \in I(\mathbf{x})$, 设 $\mathbf{d} = (d_1, \dots, d_n)^T$ 满足 $d_i \geq 0$, 并且

$$\sum_{i=1}^n d_i = 0. \text{ 若 } \mathbf{d} = \mathbf{0}, \text{ 则 } \mathbf{d} \in T_S(\mathbf{x}). \text{ 若 } \mathbf{d} \neq \mathbf{0}, \text{ 设 } \lambda = \lim_{\substack{i \notin I(\mathbf{x}) \\ d_i \neq 0}} \frac{x_i}{|d_i|}$$

> 0 , 我们有

$$x_i + \lambda d_i = \lambda d_i \geq 0, \quad i \in I(\mathbf{x}),$$

$$x_i + \lambda d_i \geq x_i - \lambda |d_i| \geq 0, \quad i \notin I(\mathbf{x}),$$

$$\sum_{i=1}^n (x_i + \lambda d_i) = 1 + 0 = 1.$$

因此, 得到 $\mathbf{x} + \lambda \mathbf{d} \in S$, 故 $\mathbf{d} \in T_S(\mathbf{x})$. 反之, 对任意的 $\mathbf{y} = (y_1, \dots, y_n)^T \in S$ 和 $\lambda > 0$, 令 $\mathbf{d} = \lambda(\mathbf{y} - \mathbf{x})$, 则有

$$d_i = \lambda(y_i - x_i) = \lambda y_i \geq 0 \quad \forall i \in I(\mathbf{x})$$

和 $\sum_{i=1}^n d_i = 0$. 因此有 $\bigcup_{\lambda > 0} \frac{1}{\lambda}(S - x) \subset \left\{ d \in T_{R_+^n}(x) \mid \sum_{i=1}^n d_i = 0 \right\}$.

注意后者是闭的, 所以 $T_S(x) = \left\{ d \in T_{R_+^n}(x) \mid \sum_{i=1}^n d_i = 0 \right\}$. \square

定理 5.1.9 设 $S_1, S_2 \subset \mathcal{B}$, $S_i \subset \mathcal{B} (i \in I, I \text{ 是指标集})$ 是凸集, $x \in \mathcal{B}$.

(1) 若 $x \in S_1 \subset S_2$, 则 $T_{S_1}(x) \subset T_{S_2}(x)$ 和 $N_{S_2}(x) \subset N_{S_1}(x)$.

(2) 若 $S = \bigcap_{i \in I(x)} S_i$ 和 $I(x) = \{i \mid x \notin \text{int} S_i\}$, 则 $T_S(x) \subset \bigcap_{i \in I(x)} T_{S_i}(x)$.

证明 由定理 5.1.1 可直接推得结论成立. \square

定理 5.1.10 设 $S_i \subset \mathcal{B} (i = 1, \dots, m)$ 是凸集, $S = \prod_{i=1}^m S_i$, $x' \in S_i (i = 1, \dots, m)$, $x = (x^1, \dots, x^m)$, 则

$$T_S(x) = \prod_{i=1}^m T_{S_i}(x^i), \quad N_S(x) = \prod_{i=1}^m N_{S_i}(x^i).$$

证明 显然有 $T_S(x) \subset \prod_{i=1}^m T_{S_i}(x^i)$. 反之, 设 $d^i \in T_{S_i}(x^i) (i = 1, \dots, m)$, 则存在 $d_k^i \rightarrow d^i (k \rightarrow \infty)$ 和 $\lambda_k^i > 0$, 使得 $x^i + \lambda_k^i d_k^i \in S_i (i = 1, \dots, m)$. 令 $\lambda = \min\{\lambda^1, \dots, \lambda^m\}$, 由于 S_i 是凸的, 我们有

$$x^i + \lambda d^i = \left(1 - \frac{\lambda}{\lambda_k^i}\right) x^i + \frac{\lambda}{\lambda_k^i} (x^i + \lambda_k^i d_k^i) \in S_i,$$

$$i = 1, \dots, m.$$

据此, 得到 $x + \lambda d \in S$, 其中 $d^i = (d_1^i, \dots, d_m^i)$, 因此 $d \in T_S(x)$.

同理, 由 (5.1.2) 可以推出第二个结论成立. \square

定理 5.1.11 设 $\varphi: \mathcal{B} \rightarrow \mathcal{B}$ 是线性映射, $S \subset \mathcal{B}$ 是凸集, $x \in S$.

(1) $T_{\varphi(S)}(\varphi(x)) = \text{cl} \varphi(T_S(x))$.

(2) $N_{\varphi(S)}(x) = \varphi^{*^{-1}}(N_S(x))$ (φ^* 是 φ 的对偶算子).

证明 因为 $\langle x^*, \varphi(x) \rangle = \max_{d \in S} \langle x^*, \varphi(d) \rangle = \max_{d \in S} \langle \varphi^*(x^*), d \rangle = \langle \varphi^*(x^*), x \rangle$, 所以(2)成立. 由(2)可以推出(1)成立. \square

推论 5.1.12 设 $S_1, S_2 \subset \mathcal{B}$ 是闭凸集, $x^1 \in S_1, x^2 \in S_2$, 则

$$T_{S_1+S_2}(x^1+x^2) = \text{cl}(T_{S_1}(x^1) + T_{S_2}(x^2)),$$

$$N_{S_1+S_2}(x^1+x^2) = N_{S_1}(x^1) \cap N_{S_2}(x^2).$$

证明 由定理 5.1.11 可以推得结论成立. \square

定理 5.1.13 设 $S \subset \mathcal{B}$ 和 $M \subset \mathcal{D}$ 是闭凸集, $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是线性映射, $V = \{x \in \mathcal{B} \mid \varphi(x) \in M\} = S \cap \varphi^{-1}(M)$. 又设 $\theta \in \varphi(S) - M$ (即 $V \neq \emptyset$), $x \in V$, 以及

$$T_V(x) \subset T_S(x) \cap \varphi^{-1}(T_M(\varphi(x))),$$

$$N_S(x) + \varphi^*(N_M(\varphi(x))) \subset N_V(x). \quad (5.1.4)$$

若 $\theta \in \text{int}(\varphi(S) - M)$, $x \in V$, 则

$$(1) T_V(x) = T_S(x) \cap \varphi^{-1}(T_M(\varphi(x)));$$

$$(2) N_V(x) = N_S(x) + \varphi^*(N_M(\varphi(x))).$$

证明 (1) 设 $d \in T_S(x) \cap \varphi^{-1}(T_M(\varphi(x)))$, 则存在点列 $\{d^k\} \subset \mathcal{B}, d^k \rightarrow d, \{y^k\} \subset \mathcal{D}, y^k \rightarrow \varphi(d)$, 以及 $\{\lambda_k^1\}$ 和 $\{\lambda_k^2\} \subset (0, +\infty), \lambda_k^1 \rightarrow 0, \lambda_k^2 \rightarrow 0 (k \rightarrow \infty)$, 使得对所有的 k , 有

$$x + \lambda_k^1 d^k \in S, \quad \varphi(x) + \lambda_k^2 y^k \in M.$$

令 $\lambda_k = \min(\lambda_k^1, \lambda_k^2)$, 因为 S 和 M 是凸的, 可以推得

$$x + \lambda_k d^k \in S, \quad \varphi(x) + \lambda_k y^k \in M, \quad k = 1, 2, \dots.$$

若 $y^k = \varphi(d^k) (k = 1, 2, \dots)$ 中有无穷多个成立, 则有 $d \in T_V(x)$.

否则, 作集值映射 $\psi: S \rightarrow 2^M, x \mapsto \psi(x)$,

$$\psi(x) = \varphi(x) - M.$$

根据已知 $\theta \in \text{int}(\psi(S)) = \text{int}(\varphi(S) - M)$, 由文献[16]的推论

3.3.4, 存在 $\delta > 0$ 使得对任意的 $y' \in \mathcal{B}$ 和 $x' \in S$, 有

$$d_{\varphi^{-1}(\mathcal{Y})}(x') \leq \frac{1}{\delta} d_{\varphi(\mathcal{Y})}(x') (\|x' - x\| + 1).$$

取 $y' = \theta$, $x' = x + \lambda_k d$, 则有 $\|x' - x\| = \lambda_k \|d\|$ 和

$$\begin{aligned} d_{\varphi(x+\lambda_k d)}(\theta) &= d_M(\varphi(x) + \lambda_k \varphi(d)) \\ &\leq d_M(\varphi(x) + \lambda_k y^k) + \lambda_k \|\varphi(d) - y^k\| \\ &= \lambda_k \|\varphi(d) - y^k\|. \end{aligned}$$

于是, 我们有

$$\begin{aligned} \frac{1}{\lambda_k} d_V(x + \lambda_k d^k) &\leq \frac{1}{\delta \lambda_k} d_{\varphi(x+\lambda_k d)}(\theta) (\|x' - x\| + 1) \\ &\leq \frac{1}{\delta} \|\varphi(x) - y^k\| (\lambda_k \|d\| + 1). \end{aligned}$$

因为 $y^k \rightarrow \varphi(d) (k \rightarrow \infty)$, 所以从上式知

$$\liminf_{\lambda_k \rightarrow \infty} \frac{1}{\lambda_k} d_V(x + \lambda_k d^k) = 0.$$

由注 5.1.1, 即得 $d \in T_V(x)$.

(2) 由(1)按(5.1.2)有

$$N_V(x) = \text{cl} \left(N_S(x) + \varphi^* \left(N_M(\varphi(x)) \right) \right).$$

现在只要证明 $N_S(x) + \varphi^* \left(N_M(\varphi(x)) \right)$ 是闭集即可. 设 $p_i^* \in N_S(x)$, $q_i^* \in N_M(\varphi(x))$, $r_k = p_i^* + \varphi^*(q_i^*)$, $r_k \rightarrow r (k \rightarrow \infty)$, 由 $\theta \in \text{int}(\varphi(S) - M)$ 可知, 存在 $\bar{\lambda} > 0$, $\bar{d} \in S$ 和 $\bar{y} \in M$, 使得对任意的 $y \in \mathcal{B}$ 有 $y = \bar{\lambda}(\bar{y} - \varphi(\bar{d}))$, 因此

$$\begin{aligned} \langle q_i^*, y \rangle &= \bar{\lambda} \langle q_i^*, \bar{y} - \varphi(\bar{d}) \rangle \\ &= \bar{\lambda} (\langle q_i^*, \bar{y} \rangle - \langle \varphi^*(q_i^*), \bar{d} \rangle) \end{aligned}$$

$$\begin{aligned}
&= \bar{\lambda}(\langle p_k^*, \bar{d} \rangle + \langle q_k^*, \bar{y} \rangle - \langle r_k, \bar{d} \rangle) \\
&\leq \bar{\lambda}(\langle p_k^*, x \rangle + \langle q_k^*, \varphi(x) \rangle - \langle r_k, \bar{d} \rangle) \\
&= \bar{\lambda}\langle r_k, x - \bar{d} \rangle \leq \bar{\lambda} \|r_k\| \cdot \|x - \bar{d}\| < +\infty.
\end{aligned}$$

根据 Banach-Steinhaus 定理(见文献[10]P. 321 的定理 1), $\{q_k^*\}$ 中存在弱收敛的子列,不妨设 $q_k^* \rightarrow q^* \in N_M(\varphi(x))$. 因此,

$$p_k^* = r_k - \varphi^*(q_k^*) \rightarrow p^* = r - \varphi^*(q^*) \in N_S(x),$$

故得 $r \in N_S(x) + \varphi^*(N_M(\varphi(x)))$. \square

由定理 5.1.13, 容易推得以下三个推论.

推论 5.1.14 设 $S \subset \mathcal{B}$ 是闭凸集, $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是线性映射, 则对任意的 $y \in \text{int} \varphi(S)$ 和 $x \in S \cap \varphi^{-1}(y)$ 有

$$T_{S \cap \varphi^{-1}(y)}(x) = T_S(x) \cap \text{Ker} \varphi,$$

$$N_{S \cap \varphi^{-1}(y)}(x) = N_S(x) + \text{Im} \varphi^*,$$

其中 $\text{Ker} \varphi$ 是 φ 的核, φ^* 是 φ 的对偶算子,

$$\text{Im} \varphi^* = \{\varphi(x^*) \mid \langle x^*, \varphi(x) \rangle = \langle \varphi^*(x^*), x \rangle, x \in S\}.$$

推论 5.1.15 设 $S \subset \mathcal{B}$ 是闭凸集, $x \in S$, $p_0^* \in S^-$, $V = \{x \in S \mid \langle p_0^*, x \rangle = -1\} \neq \emptyset$, 则 $d \in T_V(x)$ 当且仅当对任意的 $p^* \in S^-$ 满足 $\langle p^*, x \rangle = 0$, 有 $\langle p_0^*, d \rangle = 0$ 和 $\langle p^*, d \rangle \leq 0$.

推论 5.1.16 设 $S_1, S_2 \subset \mathcal{B}$ 是闭凸集. 若 $\theta \in \text{int}(S_1 - S_2)$, 则

$$T_{S_1 \cup S_2}(x) = T_{S_1}(x) \cap T_{S_2}(x) \quad \forall x \in S_1 \cap S_2.$$

再给出如下定理:

定理 5.1.17 设 $S_1, \dots, S_m \in \mathcal{B}$ 是闭凸集, $S = \bigcap_{i=1}^m S_i$. 若存在 $\bar{\delta} > 0$ 使得对任意的 $x^i \in \bar{\delta}B (i = 1, \dots, m)$ 有 $\bigcap_{i=1}^m (S_i - x^i) \neq \emptyset$, 则

$$T_S(x) = \bigcap_{i=1}^m T_{S_i}(x) \quad \forall x \in S.$$

证明 由推论 5.1.14 和定理 5.1.10 容易推得结论成立. \square

§ 5.2 D -切锥和 C -切锥

现在介绍一般集合在一点处的切锥.

设 \mathcal{B} 是 Banach 空间, \mathcal{B}^* 是 \mathcal{B} 的对偶空间, 集合 $S \subset \mathcal{B}$, 点 $x \in S$, $\varepsilon > 0$, 记 $B_S(x, \varepsilon) = S \cap (x + \varepsilon B)$, $x^k \rightarrow_S x$ 表示点列 $\{x^k\} \subset S$ 收敛于 $x \in S$.

定义 5.2.1 设集合 $S \subset \mathcal{B}$ 非空, 点 $x^0 \in \mathcal{B}$. 称集合

$$T_S^D(x^0) = \bigcap_{\varepsilon > 0} \bigcap_{\alpha > 0} \bigcup_{0 < \lambda \leq \alpha} \left(\frac{1}{\lambda} (S - x^0) + \varepsilon B \right) \quad (5.2.1)$$

是集合 S 在点 x^0 处的相依切锥, 或 D -切锥.

注 5.2.1 设 $x \in \mathcal{B}$, 由 (5.2.1) 不难得知有以下三个等价关系:

(1) $d \in T_S^D(x)$ 当且仅当对任意的 $\varepsilon > 0$ 和任意的 $\alpha > 0$, 分别存在 $x \in d + \varepsilon B$ 和 $t \in [0, \alpha]$ 使得 $x + td \in S$.

(2) $d \in T_S^D(x)$ 当且仅当存在序列 $\{d^k\} \subset \mathcal{B}$, $d^k \rightarrow d$, 以及 $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0 (k \rightarrow \infty)$, 使得对任何 $k = 1, 2, \dots$ 有 $x + t_k d^k \in S$.

(3) $d \in T_S^D(x)$ 当且仅当 $\liminf_{t \rightarrow 0^+} \frac{d_S(x + td)}{t} = 0$.

注 5.2.2 设 $x \in \mathcal{B}$, 由定义 5.2.1 易知 $T_S^D(x)$ 是闭锥. 并且, 当 $x \in \text{int} S$ 时有 $T_S^D(x) = \mathcal{B}$, $T_{\mathcal{B}}^D(x) = \mathcal{B}$ 和 $T_{\emptyset}^D(x) = \emptyset$.

设集合 $S \subset \mathcal{B}$, 点 $x \in S$, $\phi: S \rightarrow 2^{\mathcal{B}}$, $x \mapsto \phi(x)$ 是集值映射. 记

$$\liminf_{x' \rightarrow x} \phi(x') = \bigcap_{\varepsilon > 0} \bigcup_{\alpha > 0} \bigcap_{x' \in B_S(x, \alpha)} (\phi(x') + \varepsilon B),$$

它等价于:对任意的 $\varepsilon > 0$, 存在 $\eta > 0$ 使得 $\sup_{x' \in B_S(x, \eta)} d_{\varphi(x')}(d) \leq \varepsilon$.

定义 5.2.2 设集合 $S \subset \mathscr{B}$ 非空, 点 $x^0 \in \mathscr{B}$.

(1) 集合

$$\begin{aligned} T_S^C(x^0) &= \liminf_{\substack{x \rightarrow x^0 \\ \lambda \rightarrow 0^+}} \frac{1}{\lambda} (S - x) \\ &= \bigcap_{\varepsilon > 0} \bigcup_{\alpha, \beta > 0} \bigcap_{\substack{x \in B_S(x^0, \alpha) \\ \lambda \in (0, \beta)}} \left(\frac{1}{\lambda} (S - x) + \varepsilon B \right) \end{aligned} \quad (5.2.2)$$

称为集合 S 在点 x^0 处的 Clarke 切锥, 或 C-切锥.

(2) 集合

$$N_S^C(x^0) = \{x^* \in \mathscr{B}^* \mid \langle x^*, d \rangle \leq 0, d \in T_S^C(x^0)\} \quad (5.2.3)$$

称为集合 S 在点 x^0 处的 Clarke 法锥, 或 C-法锥.

注 5.2.3 设 $x \in \mathscr{B}$, 由 (5.2.2), 下面三个等价关系显然成立:

(1) $d \in T_S^C(x)$ 当且仅当对任意的 $\varepsilon > 0$, 存在 $\alpha > 0$ 和 $\beta > 0$, 使对任意的 $x' \in B_S(x, \alpha)$ 和任意的 $\lambda \in [0, \beta]$, 存在 $p \in d + \varepsilon B$ 使得 $x' + \lambda p \in S$.

(2) $d \in T_S^C(x)$ 当且仅当对任何序列 $\{x^k\} \subset S$, $x^k \rightarrow x$, $\{t_k\} \in (0, +\infty)$, $t_k \rightarrow 0 (k \rightarrow \infty)$, 存在序列 $\{d^k\} \subset \mathscr{B}$, $d^k \rightarrow d$, 使得对任何 $k = 1, 2, \dots$, 有 $x^k + t_k d^k \in S$.

(3) $d \in T_S^C(x)$ 当且仅当 $\lim_{\substack{x' \rightarrow_S x \\ t \rightarrow 0^+}} \frac{d_S(x' + td)}{t} = 0$.

注 5.2.4 显然, 当 $x \in \text{int} S$ 时有 $T_S^C(x) = \mathscr{B}$; 当 $x \in \mathscr{B}$ 时有 $T_{\mathscr{B}}^C(x) = \mathscr{B}$ 和 $T_{\emptyset}^C(x) = \emptyset$.

注 5.2.5 当 $S \subset \mathscr{B}$ 非空时, 由定理 4.4.7 和注 5.2.3 的 (2) 可知, C-切锥与 G-切锥是一致的.

定理 5.2.1 设集合 $S \subset \mathscr{B}$ 非空, $x \in \mathscr{B}$, 则 $T_S^C(x)$ 是闭

凸集.

证明 设 $d^1, d^2 \in T_S^c(x)$, 对任何收敛于 $(x, 0)$ 的序列 $\{(x^k, \lambda_k)\} \subset S \times (0, +\infty)$, 存在序列 $\{d_k^1\} \subset \mathcal{B}$ 收敛于 d^1 , 使得对任何 $k = 1, 2, \dots$, 有 $d_k = x^k + \lambda_k d_k^1 \in S$. 因为 $k \rightarrow \infty$ 时 $d_k \rightarrow x$, 由 $d^2 \in T_S^c(x)$ 知存在序列 $\{d_k^2\} \subset \mathcal{B}$ 收敛于 d^2 , 使得对任何 $k = 1, 2, \dots$, 有

$$d_k + \lambda_k d_k^2 = x^k + \lambda_k (d_k^1 + d_k^2) \in S, \quad k = 1, 2, \dots.$$

从 $d_k + d_k^2 \rightarrow d^1 + d^2$ 知 $d^1 + d^2 \in T_S^c(x)$, 因此 $T_S^c(x)$ 是闭凸集. \square

注 5.2.6 这里定义的切锥 $T_S^p(x)$ 和 $T_S^c(x)$ 与第4章中引进的 G -切锥有如下关系:

$$T_S^c(x) \subset T_S^p(x) \subset T_S^g(x). \quad (5.2.4)$$

当 S 是凸集时, (5.2.4) 的包含关系成为等式.

定理 5.2.2 设集合 $S \subset \mathcal{B}$ 非空, $x \in \mathcal{B}$. 若 S 是凸集, 则

$$T_S^c(x) = T_S^p(x) = T_S^g(x) = T_S(x). \quad (5.2.5)$$

证明 由定理 4.4.5 和 (5.2.4), 只需证明 $T_S(x) \subset T_S^c(x)$.

设 $d \in T_S(x)$, 由 (5.1.1) 有 $d \in \text{cl} \bigcup_{\lambda > 0} \frac{1}{\lambda}(S - x)$. 取任给的 $\varepsilon > 0$, 则存在 $y \in S$ 和 $\beta > 0$, 使得 $d - \frac{1}{\beta}(y - x) \in (\varepsilon/2)B$. 令 $\alpha = (\varepsilon/2)\beta$, $x' \in B_S(x, \alpha)$, $\lambda \in (0, \beta)$ 和 $p = (y - x')/\beta$, 则有

$$x + \lambda p = (1 - \lambda/\beta)x + (\lambda/\beta)y \in S,$$

$$\|p - d\| \leq (1/\beta)\|x' - x\| + \|d - (y - x')/\beta\|$$

$$\leq \alpha/\beta + \varepsilon/2 = \varepsilon.$$

因此, 按定义 5.2.2 的 (1) 得 $d \in T_S^c(x)$. 定理得证. \square

通过下面的例子可以看到 $T_S^c(x)$ 和 $T_S^p(x)$ 不一定相等.

例 5.2.1 考虑函数 $f: \mathbb{R} \rightarrow \mathbb{R}$, 当 $x \leq 0$ 时 $f(x) = 0$, 当 $x \geq$

0 时 $f(x) = x$. 取 $S = \{(x, f(x)) | x \in R\}$.

若 $x < 0$, 则 $T_S^C(x, 0) = T_S^D(x, 0) = R \times \{0\}$.

若 $x = 0$, 则 $T_S^C(0, 0) = (0, 0)$,

$$T_S^D(0, 0) = (-R_+ \times \{0\}) \cup \{(x', x') | x' \geq 0\}.$$

若 $x > 0$, 则 $T_S^C(x, x) = T_S^D(x, x) = \{(x', x') | x' \in R\}$.

定理 5.2.3 设集合 $S \subset R^n$ 非空, 则对任意的 $x \in S$, 有

$$T_S^C(x) \subset \liminf_{x' \rightarrow x} T_S^D(x').$$

证明 由定义 5.2.2, 有

$$T_S^C(x) = \bigcap_{\epsilon > 0} \bigcup_{\alpha, \beta > 0} \bigcap_{x' \in B_S(x, \alpha)} \bigcap_{\lambda \in (0, \beta)} \left(\frac{1}{\lambda} (S - x') + \epsilon B \right).$$

固定 ϵ 和 α , 则

$$\begin{aligned} & \bigcup_{\beta > 0} \bigcap_{x' \in B_S(x, \alpha)} \bigcap_{\lambda \in (0, \beta)} \left(\frac{1}{\lambda} (S - x') + \epsilon B \right) \\ & \subset \bigcap_{x' \in B_S(x, \alpha)} \bigcup_{\beta > 0} \bigcap_{\lambda \in (0, \beta)} \left(\frac{1}{\lambda} (S - x') + \epsilon B \right). \end{aligned}$$

因为 S 是有限维的, 故有

$$\bigcup_{\beta > 0} \bigcap_{\lambda \in (0, \beta)} \left(\frac{1}{\lambda} (S - x') + \epsilon B \right) \subset T_S^D(x') + \epsilon B.$$

据此, 由定义 5.2.1 知

$$T_S^C(x) \subset \bigcap_{\epsilon > 0} \bigcup_{\alpha > 0} \bigcap_{x' \in B_S(x, \alpha)} (T_S^D(x') + \epsilon B) = \liminf_{x' \rightarrow_S x} T_S^D(x'),$$

于是结论得证. \square

定理 5.2.4 设集合 $S_i \subset \mathcal{B}$ 非空 ($i \in I = \{1, 2, \dots\}$).

(1) 若 $S_1 \subset S_2$, 则对任意的 $x \in S_1$, 有 $T_{S_1}^D(x) \subset T_{S_2}^D(x)$.

(2) 若 $S = \bigcup_{i \in I} S_i$, 则

$$\bigcup_{i \in I} T_{S_i}^D(x) \subset T_S^D(x) \quad \forall x \in S.$$

若 I 是有限集, 则上式成为等式.

(3) 若 $S = \bigcap_{i \in I} S_i$, 则

$$T_S^D(x) \subset \bigcap_{i \in I} T_{S_i}^D(x) \quad \forall x \in S.$$

(4) 若 $I = \{1, \dots, m\}$, $S = \prod_{i \in I} S_i$, 则对任意的 $x = (x^1, \dots, x^m) \in S$, 有

$$T_S^D(x) \subset \prod_{i \in I} T_{S_i}^D(x^i), \quad T_S^C(x) = \prod_{i \in I} T_{S_i}^C(x^i).$$

证明 由定义 5.2.1 和定义 5.2.2 容易推得结论. \square

设 \mathscr{D} 是 Banach 空间.

定理 5.2.5 设 $\Omega \subset \mathscr{B}$ 是开集. 若映射 $\varphi: \Omega \rightarrow \mathscr{D}$ 是严格可微的, 集合 $S \subset \Omega$, 则对任意的 $x \in S$ 有

$$\langle D_S \varphi(x), T_S^D(x) \rangle \subset T_{\varphi(S)}^D(\varphi(x)).$$

证明 设 $d \in T_S^D(x)$, 根据注 5.2.1 的(2), 存在序列 $\{d^k\} \subset \mathscr{B}$, $d^k \rightarrow d$, $\{t_k\} \subset (0, +\infty)$, $t_k \rightarrow 0^+$ ($k \rightarrow \infty$), 使得对任何 $k = 1, 2, \dots$, 有 $x + \lambda_k d^k \in S$. 由 φ 是严格可微的, 根据定义 4.2.3 有

$$x^k = \frac{\varphi(x + t_k d^k) - \varphi(x)}{t_k} \rightarrow \langle D_S \varphi(x), d \rangle,$$

从而 $\varphi(x) + \lambda_k x^k = \varphi(x + t_k d^k) \in \varphi(S)$. 由此, 得到 $\langle D_S \varphi(x), d \rangle \in T_{\varphi(S)}^D(\varphi(x))$. \square

注 5.2.7 特别地, 当 $\varphi: \mathscr{B} \rightarrow \mathscr{D}$ 是线性映射时, 我们有 $\varphi(T_S^D(x)) \subset T_{\varphi(S)}^D(\varphi(x))$.

定理 5.2.6 设集合 $S \subset \mathscr{B}$, $M \subset \mathscr{D}$. 若 φ 是 S 的一个开邻域到 \mathscr{D} 的严格可微映射,

$$V = \{x \in S \mid \varphi(x) \in M\} = S \cap \varphi^{-1}(M),$$

则

$$T_V^D(\mathbf{x}) \subset T_S^D(\mathbf{x}) \cap D_S\varphi(\mathbf{x})^{-1}T_M^D(\varphi(\mathbf{x})).$$

证明 由定理 5.2.5 有 $T_V^D(\mathbf{x}) \subset T_S^D(\mathbf{x})$. 再据定理 5.2.6 得

$$\langle D_S\varphi(\mathbf{x}), T_V^D(\mathbf{x}) \rangle \subset T_{\varphi(S)}^D(\varphi(\mathbf{x})) \subset T_M^D(\varphi(\mathbf{x})).$$

因此有 $T_V^D(\mathbf{x}) \subset D_S\varphi(\mathbf{x})^{-1}T_M^D(\varphi(\mathbf{x}))$. \square

定理 5.2.7 设集合 $S \subset \Omega \subset R^n$, $M \subset R^m$, 并且 Ω 是开集, S 和 M 是闭集. 又设映射 $\varphi: \Omega \rightarrow R^m$ 是严格可微的, $V = \{\mathbf{x} \in S \mid \varphi(\mathbf{x}) \in M\} = S \cap \varphi^{-1}(M)$. 若存在 $\mathbf{x} \in V$, 使得

$$\langle D_S\varphi(\mathbf{x}), T_S^C(\mathbf{x}) \rangle - T_M^D(\varphi(\mathbf{x})) = R^m,$$

则

$$T_S^C(\mathbf{x}) \cap D_S\varphi(\mathbf{x})^{-1}T_M^C(\varphi(\mathbf{x})) \subset T_V^C(\mathbf{x}).$$

证明 作集值映射 $\psi: R^n \rightarrow 2^{R^m}$, $\mathbf{x} \mapsto \psi(\mathbf{x})$,

$$\psi(\mathbf{x}) = \begin{cases} \varphi(\mathbf{x}) - M, & \mathbf{x} \in S; \\ \emptyset, & \mathbf{x} \notin S. \end{cases}$$

显然, 有 $\psi^{-1}(\mathbf{0}) = V$. 已知 $\mathbf{x} \in V$, 则存在 \mathbf{x} 的一个开邻域 $U(\mathbf{x})$ 以及 $\alpha > 0$ 和 $\beta > 0$, 使得

$$d_{\psi^{-1}(\mathbf{0})}(\mathbf{x}') \leq \beta \| \mathbf{y} \| \quad \forall \mathbf{y} \in \alpha B, \quad \forall \mathbf{x}' \in \psi^{-1}(\mathbf{y}) \cap U(\mathbf{x}).$$

选择充分小的 α , 使得

$$\| \varphi(\mathbf{x}') - \varphi(\mathbf{x}) \| < \alpha,$$

$$S \cap U(\mathbf{x}) \subset \psi^{-1}(\varphi(\mathbf{x}') - \pi_M(\varphi(\mathbf{x}'))),$$

其中 $\pi_M(\varphi(\mathbf{x}')) = \{\varphi(\mathbf{x}') \in M \mid \|\varphi(\mathbf{x}') - \mathbf{y}\| = d_M(\mathbf{y}), \mathbf{x}' \in R^n\}$,

则 $\|\varphi(\mathbf{x}') - \pi_M(\varphi(\mathbf{x}'))\| \leq \|\varphi(\mathbf{x}') - \varphi(\mathbf{x})\| < \alpha$.

由此, 对任意的 $\mathbf{x}' \in S \cap U(\mathbf{x})$ 有

$$d_{\psi^{-1}(\mathbf{0})}(\mathbf{x}') \leq \beta \|\varphi(\mathbf{x}') - \pi_M(\varphi(\mathbf{x}'))\| = \beta d_M(\varphi(\mathbf{x}')).$$

(5.2.6)

下面设 $d \in T_S^c(x) \cap D_S\varphi(x)^{-1}T_M^c(\varphi(x))$, 则存在 $\varepsilon > 0$, $\delta > 0$ 使得当 $\|x' - x\| \leq \varepsilon$ 和 $\lambda \leq \delta$ 时, 有 $x + \lambda d \in S \cap U(x)$. 从 (5.2.6) 我们有

$$\begin{aligned} \frac{1}{\lambda} d_{\varphi^{-1}(U)}(x' + \lambda d) &\leq \beta \frac{d_M(\varphi(x' + \lambda d))}{\lambda} \\ &\leq \beta \frac{d_S(\varphi(x')) + \lambda \langle D_S\varphi(x), d \rangle}{\lambda} \\ &\quad + \beta \frac{\|\varphi(x' + \lambda d) - \varphi(x') - \lambda \langle D_S\varphi(x), d \rangle\|}{\lambda}. \end{aligned}$$

因为 $\langle D_S\varphi(x), d \rangle \in T_M^c(\varphi(x))$, 所以上式右端两式当 $\lambda \rightarrow 0$ 时趋于 0, 从而得 $d \in T_V^c(x)$. \square

由定理 5.2.7, 容易推知以下三个推论成立.

推论 5.2.8 设 $\varphi: R^n \rightarrow R^m$ 是连续严格可微映射, $M \subset \mathscr{D}$ 是闭集, $\varphi(x) \in M$. 若 $\text{Im} D_S\varphi(x) = T_M^c(\varphi(x)) = M$, 则

$$D_S\varphi(x)^{-1}T_M^c(\varphi(x)) \subset T_{\varphi^{-1}(M)}^c(x).$$

推论 5.2.9 设 $S_1, S_2 \subset R^n$ 是非空闭集, $x \in S_1 \cap S_2$. 若 $T_{S_1}^c(x) = T_{S_2}^c(x) = R^n$, 则

$$T_{S_1}^c(x) \cap T_{S_2}^c(x) \subset T_{S_1 \cup S_2}^c(x).$$

推论 5.2.10 设 $S_i \subset R^n$ 是非空闭集 ($i = 1, \dots, m$), $x \in \bigcap_{i=1}^m S_i$. 若对任意的 $d^1, \dots, d^m \in R^n$ 有 $\bigcap_{i=1}^m (T_{S_i}^c(x) - d^i) \neq \emptyset$, 则

$$\bigcap_{i=1}^m T_{S_i}^c(x) \subset T_{\bigcap_{i=1}^m S_i}^c(x).$$

§ 5.3 D -切导数

从这一节开始, 我们将利用切锥来定义导数和次微分. 设 \mathscr{B}

是 Banach 空间, $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数. 作集值映射 $f_+: \mathcal{B} \rightarrow 2^R, x \mapsto R,$

$$f_+(x) = \begin{cases} f(x) + R_+, & f(x) < +\infty; \\ \emptyset, & f(x) = +\infty. \end{cases}$$

定义 5.3.1 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 点 $x^0 \in \mathcal{B}, d \in \mathcal{B}, T_{\text{epi}f_+}^D(x^0, f(x^0))$ 是 $\text{epi}f_+$ 在点 $(x^0, f(x^0))$ 处的 D -切锥. 又设 $(d, \eta) \in T_{\text{epi}f_+}^D(x^0, f(x^0))$, 记

$$f^D(x^0; d) = \inf \{ \eta \mid (d, \eta) \in T_{\text{epi}f_+}^D(x^0, f(x^0)) \}, \quad (5.3.1)$$

则称 $f^D(x^0; d)$ 是函数 f 在点 x^0 处沿方向 d 的 D -切导数.

注 5.3.1 在 $T_{\text{epi}f_+}^D(x, f(x))$ 中对每一给定的 (d, η) 构成的集合或为 R , 或为 $(\eta, +\infty)$, 或为空集.

由于 D -切锥不一定是凸的, 故不用它来定义相应的次梯度. 因为即使给出次梯度的定义, 也得不到多少好的性质. 下面研究 D -切导数的一些性质.

定理 5.3.1 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 则

$$f^D(x; d) = \liminf_{\substack{d' \rightarrow d \\ t \rightarrow 0^+}} \frac{f(x + td') - f(x)}{t}.$$

并且, 若对任意的 $d \in \mathcal{B}$ 有 $f^D(x; d) < +\infty$, 则 $f^D(x; d)$ 关于 d 是正齐次和下半连续的.

证明 设 $(d, \eta) \in T_{\text{epi}f_+}^D(x, f(x))$, 则对任意的 $\epsilon_1 > 0$ 和 $\epsilon_2 > 0$, 以及任意的 $\alpha > 0$, 存在 $d' \in d + \epsilon_2 B$ 和 $t \in [0, \alpha]$, 使得

$$\eta \in \frac{f_+(x + td') - f(x)}{t} + \epsilon_1 B',$$

其中 $B' = [-1, 1]$. 从上式可推得

$$\epsilon_1 + \eta \geq \frac{f(x + td') - f(x)}{t}$$

$$\geq \inf_{t < \alpha} \inf_{\|d' - d\| \leq \varepsilon_2} \frac{f(x + td') - f(x)}{t},$$

因此

$$\eta \geq \liminf_{\substack{d' \rightarrow d \\ t \rightarrow 0^+}} \frac{f(x + td') - f(x)}{t} = \varepsilon_1.$$

设 $\alpha = \liminf_{\substack{d' \rightarrow d \\ t \rightarrow 0^+}} \frac{f(x + td') - f(x)}{t} = \varepsilon_1$, 则有 $\alpha \leq f^D(x; d)$.

另一方面, 对任意的 $M > \alpha$, 有

$$\sup_{\substack{\alpha > 0 \\ \delta > 0}} \inf_{t < \alpha} \inf_{\|d' - d\| \leq \delta} \frac{f(x + td') - f(x)}{t} < M,$$

即对任意的 $\alpha > 0$, $\delta > 0$, 存在 $t < \alpha$ 和 $d' \in d + \delta B$, 使得

$$\frac{f(x + td') - f(x)}{t} \leq M.$$

由此, 得 $M \in \frac{f(x + td') - f(x)}{t}$, 也就是说 $(d, a) \in T_{\text{epif}_+}^D(x, f(x))$. 再由定义 5.3.1 便知 $a = f^D(x; d)$. 其余结论容易推得. \square

注 5.3.2 由定理 5.3.1 可知, 若 $f: R^n \rightarrow R$ 在 $x \in R^n$ 处可微, 则

$$f^D(x; d) = \langle \nabla f(x), d \rangle.$$

若 f 是凸函数, 则对任意的 $d \in R^n$ 有

$$f^D(x; d) = \liminf_{d' \rightarrow d} \left(\inf_{t > 0} \frac{f(x + td') - f(x)}{t} \right).$$

定理 5.3.2 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数. 若 f 在 $x \in \text{int}(\text{dom} f)$ 处是 Lipschitz 的, 则

$$f^D(x; d) = \liminf_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} \quad \forall d \in \mathcal{B},$$

并且 $f^D(x; d)$ 有限.

证明 由定理 5.3.1 的证明, 可知结论成立. \square

定理 5.3.3 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 集合 $S \subset \mathcal{B}$ 非空, 记

$$f|_S(x) = \begin{cases} f(x), & x \in S; \\ +\infty & x \notin S, \end{cases}$$

则

$$f^D(x; d) \leqslant (f|_S)^D(x; d) \quad \forall x \in \mathcal{B}, \forall d \in T_S^D(x).$$

若 f 在 $x \in S$ 处是严格可微的, 则有

$$(f|_S)^D(x; d) = \begin{cases} \langle D_S f(x), d \rangle, & d \in T_S^D(x); \\ +\infty, & d \notin T_S^D(x). \end{cases}$$

证明 由定义 5.3.1, 易知结论成立. \square

定理 5.3.4 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 若 $x^0 \in \text{dom} f$, 并且对任意的 $x \in \mathcal{B}$ 有 $f(x^0) \leqslant f(x)$, 则

$$f^D(x; d) \geqslant 0 \quad \forall d \in \mathcal{B}.$$

证明 由定理 5.3.1 直接可得. \square

以下为 ϵ -变分原理.

定理 5.3.5 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是下半连续和有界的正常函数, $x^0 \in \text{dom} f$, 则对任意的 $\epsilon > 0$ 存在 $x_\epsilon \in \text{dom} f$, 使得

$$f(x_\epsilon) + \epsilon \|x_\epsilon - x^0\| \leqslant f(x^0), \quad (5.3.2)$$

$$0 \leqslant f^D(x_\epsilon; d) + \epsilon \|d\| \quad \forall d \in \mathcal{B}. \quad (5.3.3)$$

证明 根据文献[16]的定理 5.3.1, 得知存在 $x_\epsilon \in \text{dom} f$ 满足(5.3.2), 并且

$$f(x_\epsilon) = \min \{f(x) + \epsilon \|x_\epsilon - x\| \mid x \in \mathcal{B}\}.$$

设 $d \in \text{dom}(f^D(x_\epsilon; \cdot))$, 则对任意的 $\eta > 0, \delta > 0, \alpha > 0$, 存在 $t \leqslant \alpha$ 和 $d' \in d + \delta B$, 使得

$$\frac{f(x_i + td') - f(x_i)}{t} \leq f^D(x_i; d) + \eta.$$

再由文献[16]的定理 5.3.1 可得

$$-\varepsilon\delta - \varepsilon\|d\| \leq -\varepsilon\|d\| \leq \frac{f(x_i + td') - f(x_i)}{t},$$

因此有

$$0 \leq f^D(x_i; d) + \varepsilon\|x\| + \varepsilon\delta + \eta.$$

令 $\delta, \eta \rightarrow 0$, 即得 (5.3.3) 成立. \square

定理 5.3.6 设 $f_1, f_2: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 则

$$f_1^D(x; d) + f_2^D(x; d) \leq (f_1 + f_2)^D(x; d).$$

证明 由定义 5.3.1 直接可得到. \square

定理 5.3.7 设 \mathcal{D} 是 Banach 空间, $f: \mathcal{D} \rightarrow R \cup \{+\infty\}$ 是正常函数, $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是严格可微映射. 若 $\varphi(x) \in \text{dom} f$, 则

$$f^D(\varphi(x); \langle D_\varphi \varphi(x), d \rangle) \leq (f(\varphi))^D(x; d) \quad \forall d \in \mathcal{B}.$$

证明 由定理 5.2.6 可得到. \square

§ 5.4 C-切导数和 C-切微分

本节利用函数上图象的 C-切锥来定义导数. 由于 C-切锥是凸的, 因而还可以进而引入相应的次梯度和次微分. 我们将看到, 当函数满足 Lipschitz 条件时, 本节定义的导数与第 4 章引进的 G-方向导数是一致的.

设 \mathcal{B} 是 Banach 空间.

定义 5.4.1 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 点 $x^0 \in \mathcal{B}$, $d \in \mathcal{B}$, $T_{\text{epi} f_+}^C(x^0, f(x^0))$ 是 $\text{epi} f_+$ 在点 $(x^0, f(x^0))$ 处的 C-切锥. 又设 $(d, \eta) \in T_{\text{epi} f_+}^C(x^0, f(x^0))$, 记

$$f^c(x^0; d) = \inf \{ \eta \mid (d, \eta) \in T_{\text{epi}f_+}(x^0, f(x^0)) \}, \quad (5.4.1)$$

则称 $f^c(x^0; d)$ 是 f 在点 x^0 处沿 d 的上图象导数, 或 C-切导数.

定理 5.4.1 设 $x \in \mathcal{B}$. 若对任意的 $d \in \mathcal{B}$, $f^c(x; d) > -\infty$, 则 $f^c(x; d)$ 关于 d 是正齐次和下半连续的凸函数.

证明 由 C-切锥的定义和性质即可得到. \square

设 \mathcal{B}^* 是 \mathcal{B} 的对偶空间.

定义 5.4.2 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, 点 $x^0 \in \mathcal{B}$, 则称集合

$$\partial^c f(x^0) = \{x^* \in \mathcal{B}^* \mid \langle x^*, d \rangle \leq f^c(x^0; d) \quad \forall d \in \mathcal{B}\} \quad (5.4.2)$$

是 f 在点 x^0 处的 C-切微分.

注 5.4.1 若存在 $d \in \mathcal{B}$ 使得 $f^c(x; d) = -\infty$, 则 $\partial^c f(x) = \emptyset$. 若 f 在点 x 处严格可微, 则 $f^c(x; d) = \langle D_s f(x), d \rangle$, $\partial^c f(x) = \{D_s f(x)\}$. 当 f 是凸函数时, 这里定义的 C-次微分与第 3 章定义的次微分是一致的.

定理 5.4.2 设 $f: \mathcal{B} \rightarrow R$ 是实值函数, $x \in \text{int}(\text{dom} f)$. 若 f 在 x 处是 Lipschitz 的, 则

$$f^c(x; d) = f^0(x; d).$$

证明 由定理 4.4.14 可以推得. \square

设 $x, x' \in \mathcal{B}$, $\alpha \in R$, 记号 $(x', \alpha) \downarrow x$ 表示 $\alpha \geq f(x')$, $x' \rightarrow x$, $\alpha \rightarrow f(x)$.

定理 5.4.3 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数. 若 $x \in \text{dom} f$, $\alpha \in R$, 则

$$f^c(x; d) = \lim_{\substack{(x', \alpha) \downarrow x \\ t \rightarrow 0^+}} \sup_{d' \rightarrow d} \inf \frac{f(x' + td') - \alpha}{t}$$

即

$$f^c(x; d) = \lim_{\substack{\varepsilon \rightarrow 0^+, (x', \alpha) \downarrow x \\ t \rightarrow 0^+}} \sup_{d' \in d + \varepsilon B} \inf_{t} \frac{f(x' + td') - \alpha}{t}.$$

证明 取 $(d, \eta) \in T_{\text{epi}f_+}^c(x, f(x))$, 任给 $\varepsilon > 0$, 设 $(x^k, \alpha_k) \downarrow x$, $t_k \rightarrow 0^+$ ($k \rightarrow \infty$), 则存在 $(d^k, \eta_k) \rightarrow (d, \eta)$ 使得 $(x^k, \alpha_k) + t_k(d^k, \eta_k) \in \text{epi}f_+$, 即

$$\alpha_k + t_k \eta_k \geq f(x^k + t_k d^k).$$

由此, 我们有

$$\inf_{d' \rightarrow d} \frac{f(x^k + t_k d') - \alpha_k}{t} \leq \frac{f(x^k + t_k d^k) - \alpha_k}{t} \leq \eta_k,$$

从而得 $f^c(x; d) \leq \eta$.

反之, 设 $f^c(x; d) \leq \eta$, $(x^k, \alpha_k) \downarrow x$, $t_k \rightarrow 0^+$ ($k \rightarrow \infty$), 则对每一个 $k = 1, 2, \dots$, 存在一个 $\varepsilon_k < 1/k$, 使得

$$\lim_{\substack{(x^k, \alpha) \downarrow x \\ t_k \rightarrow 0^+}} \sup_{d' \rightarrow d} \inf_{t_k} \frac{f(x^k + t_k d') - \alpha}{t_k} \leq \eta + \frac{1}{k}.$$

从上式有

$$\inf_{d' \in d + \varepsilon_k B} \frac{f(x^k + t_k d') - \alpha_k}{t_k} \leq \eta + \frac{2}{k},$$

故对每一 ε_k , 存在 $d^k \in d + \varepsilon_k B$ 使得

$$\frac{f(x^k + t_k d^k) - \alpha_k}{t_k} \leq \eta + \frac{3}{k}.$$

由此, 定义 $\eta_k = \max \left\{ \eta, \frac{f(x^k + t_k d^k) - \alpha_k}{t_k} \right\}$ ($k = 1, 2, \dots$), 有 $\eta_k \rightarrow \eta$ ($k \rightarrow \infty$), 故

$$\alpha_k + t_k \eta_k \geq f(x^k + t_k d^k), \quad k = 1, 2, \dots.$$

由定义 5.2.2 知 $(d, \eta) \in T_{\text{epi}f_+}^c(x, f(x))$, 再据定义 5.4.1 便知

结论成立. \square

推论 5.4.4 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数. 若 $x \in \text{dom} f$, 则 $\partial^c f(x) = \emptyset$ 当且仅当 $f^c(x; d) = +\infty$. 否则, 有

$$\partial^c f(x) = \{x^* \in \mathcal{B}^* \mid \langle x^*, d \rangle \leq f^c(x; d) \quad \forall d \in \mathcal{B}\}$$

和

$$f^c(x; d) = \sup \{\langle x^*, d \rangle \mid x^* \in \partial^c f(x)\}.$$

证明 由定义 5.4.1、定义 5.4.2 和定理 5.4.3 可得到. \square

§ 5.5 H -切导数

这一节, 通过 § 4.4 中的 H -切锥来引进另一种切导数. 它与 C -切导数有着紧密的联系. 在 § 5.6 中将再进一步讨论 C -切导数的有关问题.

定义 5.5.1 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是实值函数, 点 $x^0 \in \mathcal{B}$, $d \in \mathcal{B}$, $T_{\text{epi} f}^H(x^0, f(x^0))$ 是 $\text{epi} f_+$ 在点 $(x^0, f(x^0))$ 处的 H -切锥. 称

$$f''(x; d) = \inf \{\eta \mid (d, \eta) \in T_{\text{epi} f}^H(x^0, f(x^0))\} \quad (5.5.1)$$

是函数 f 在点 x^0 处沿方向 d 的 H -切导数. 若 $f''(x^0; d) < +\infty$, 则称 f 在点 x^0 处关于 d 是有限 H -可切导的. 设 $x \in \mathcal{B}$, 若函数 $f'': \mathcal{B} \times \mathcal{B} \rightarrow R$, $(x, d) \mapsto f''(x; d)$ 关于 d 是正常的 (即存在 \bar{d} 使得 $f''(x; \bar{d}) < +\infty$), 则称 f 在点 x 处关于 d 是正常 H -可切导的.

注 5.5.1 显然, 若 f 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, 则 f 在 x 处关于任意的 $d \in \mathcal{B}$ 是正常 H -可切导的. 若 $d \in \text{dom} f''(x; \cdot)$, 则 f 在 x 处关于 d 是有限 H -可切导的.

注 5.5.2 设 $x \in \mathcal{B}$, 由定义 5.3.1、定义 5.4.1 和定义 5.5.1, 可推知有

$$f^D(x; d) \leq f^C(x; d) \leq f^H(x; d) \quad \forall d \in \mathcal{B}.$$

定理 5.5.1 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是实值函数, $x \in \text{dom} f$. 若 $f^H(x; d) < +\infty$, 则

$$f^H(x; d) = \lim_{\substack{(x', \alpha) \downarrow x, d' \rightarrow d \\ t \rightarrow 0^+}} \sup \frac{f(x' + td') - \alpha}{t}.$$

证明 设 $(d, \eta) \in T_{\text{epi} f_+}^H(x, f(x))$, 根据定义 4.4.3, 存在 $\epsilon > 0$ 有

$$\begin{aligned} & ((x, f(x)) + \epsilon B_{\mathcal{B} \times R}) \cap (\text{epi} f_+ + t((d, \eta) + \epsilon B_{\mathcal{B} \times R})) \\ & \subset \text{epi} f_+, \quad t \in (0, \epsilon). \end{aligned} \quad (5.5.2)$$

设 $(x', \alpha) \downarrow x, d' \rightarrow d$, 从上式得 $\alpha + t\eta \geq f(x' + td')$, 即

$$\frac{f(x' + td') - \alpha}{t} \leq \eta.$$

因此, 由 (5.5.1) 有 $f^H(x; d) \leq \eta$.

反之, 设 $f^H(x; d) \leq \eta$, 由定义 5.5.1 可知, 对所有的 $(x', \alpha) \downarrow x, d' \rightarrow d, t \rightarrow 0^+$, 有

$$\frac{f(x' + td') - \alpha}{t} \leq \eta.$$

这说明存在 $\epsilon > 0$ 使得 (5.5.2) 成立, 于是 $(d, \eta) \in T_{\text{epi} f_+}^H(x, f(x))$. \square

定理 5.5.2 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数. 若下面 5 个条件之一成立, 则对任意的 $d \in \mathcal{B}$, f 在 x 处关于 d 是有限 H -可切导的:

- (1) f 在 x 处是 Lipschitz 的.
- (2) $f = \delta_{\text{epi} f_+}(x)$, 并且 $\text{epi} f_+$ 在 x 处存在超切向量.
- (3) f 是凸函数, 并且在某点 (不必是 x) 的邻域有界.

(4) f 关于由某个内部非空的闭凸锥 $K \in \mathcal{B}$ 确定的偏序是非减函数 (即由 $x_2 - x_1 \in K$, 有 $f(x_1) \leq f(x_2)$).

(5) $\mathcal{B} = R^n$, f 在 x 的一个邻域内是上半连续的, 并且 $\partial^c f(x) \neq \emptyset$.

证明 (1)(2) 由定理 5.5.1 即可得到.

(3) 记 $C = \text{epi} f_+$. 设 f 在 $x + t\mathbf{d}$ 的某个邻域 U 内有界, 则对某个 $t_0 > 0$, 存在 $\beta \in (1, +\infty)$ 有

$$|f(x')| < \beta - 1 \quad \forall x' \in U.$$

因此, 有 (\mathbf{d}, α) 使得 $(x, f(x)) + t_0(\mathbf{d}, \alpha) \in \text{int} C$, 其中 $\alpha = \frac{\beta - f(x)}{t_0}$. 记 $B = B_{\mathcal{B}} \times B_R$, 可以选择 $\delta > 0$, 使得

$$((x, f(x)) + \delta B) + t_0((\mathbf{d}, \alpha) + \delta B) \subset C.$$

已知 f 是凸的, 容易验证 C 也是凸的. 由上式知对任意的 $t \in (0, t_0)$, $x' \in x + \delta B_{\mathcal{B}}$, $\mathbf{d}' \in \mathbf{d} + \delta B_{\mathcal{B}}$ 和 $f(x') \in f(x) + \delta B_R$, 有 $r', r'' \in \alpha + \delta B_R$ 使得 $f(x') \leq r'$ 和 $f(x' + t_0 \mathbf{d}') \leq r''$ 成立, 并且有

$$\begin{aligned} x' + tx &= \left(1 - \frac{t}{t_0}\right)x' + \frac{t}{t_0}(x' + t_0 \mathbf{d}'), \\ f(x' + t\mathbf{d}') &\leq \left(1 - \frac{t}{t_0}\right)f(x') + \frac{t}{t_0}f(x' + t_0 \mathbf{d}') \\ &\leq \left(1 - \frac{t}{t_0}\right)r' + \frac{t}{t_0}(r' + t_0 r'') \leq r' + tr''. \end{aligned}$$

据此, 得到

$$\begin{aligned} ((x, f(x)) + \delta B) \cap C + t[(\mathbf{d}, \alpha) + \delta B] &\subset C \\ \forall t \in (0, t_0), \end{aligned}$$

故有 $(\mathbf{d}, \alpha) \in T_{\text{epi} f_+}^H(x, f(x))$. 由定理 5.5.1, 即知 $f^H(x; \mathbf{d})$ 存在.

(4) 从已知可设 $-\mathbf{d} \in \text{int} K$ 使得 $-\mathbf{d} + \varepsilon B \subset K$ (其中 $\varepsilon > 0$). 于是, 对 $-p \in -\mathbf{d} + \varepsilon B$ 和 $t > 0$ 有 $-tp \in K$. 因为 f 关于

K 是非减的, 故对任意的 y , 由 $y - (y + tp) \in K$ 有 $f(y + tp) \leq f(y)$, 从而得

$$\frac{f(y + tp) - f(y)}{t} \leq 0.$$

因此, 当 $y \rightarrow x$, $p \rightarrow d$ 和 $t \rightarrow 0^+$ 时, 上式有界, 也即 $f^H(x; d) < +\infty$.

(5) 由 $\partial^C f(x) \neq \emptyset$, 令 $D = \{d \mid f^C(x; d) < +\infty\}$, 从推论 5.4.4 知 $\text{int} D \neq \emptyset$. 因为 \mathcal{B} 是有限维的, 并且 $f^C(x; d)$ 在 $\text{int} D$ 内是有界的, 利用推论 4.5.11, 即得 $(d, \eta) \in \text{int}(\text{epi} f^C(x; d)) = \text{int} T_{\text{epi} f_-}^C(x, f(x)) = T_{\text{epi} f_+}^H(x, f(x))$. \square

§ 5.6 若 干 关 系

现在继续讨论 C -切导数和 C -切微分的性质, 也讨论与 D -切导数有关的性质.

设 \mathcal{B} 是 Banach 空间, \mathcal{B}^* 是 \mathcal{B} 的对偶空间, 记

$$\text{dom} f^C(x; \cdot) = \{d \in \mathcal{B} \mid f^C(x; \cdot) < +\infty\},$$

$$\text{dom} f^H(x; \cdot) = \{d \in \mathcal{B} \mid f^H(x; \cdot) < +\infty\}.$$

定理 5.6.1 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, $x \in \text{dom} f$, $d \in \mathcal{B}$, f 在 x 处关于 d 是正常 H -可切导的.

$$(1) \text{int dom} f^C(x; \cdot) = \text{dom} f^H(x; \cdot).$$

$$(2) f^C(x; d) = \liminf_{d' \rightarrow d} f^H(x; d') \quad (\forall d \in \text{dom} f^C(x; \cdot)).$$

(3) f^C 在 $(x, d') \mid d' \in \text{int dom} f^C(x; \cdot)$ 处是上半连续的.

(4) f^C 关于 d 在 $d' \in \text{int dom} f^C(x, \cdot)$ 处是连续的.

(5) 若 $\text{dom} f^H(x, \cdot) = \mathcal{B}$, 则 $\partial^C(-f)(x) = -\partial^C f(x)$.

证明 (1) 设 $d \in \text{dom} f^H(x; \cdot)$, 由注 5.5.2 有 $f^C(x; d) \leq$

$f^H(x; d)$, 所以 $d \in \text{int dom}(f^c(x; \cdot))$. 反之, 设 $d \in \text{int dom } f^c(x; \cdot)$, 根据定义 5.5.1 有

$$f^H(x; d' + d'') \leq f^H(x; d') + f^c(x; d'') \quad \forall d', d'' \in \mathcal{B}. \quad (5.6.1)$$

于是, 存在 $\lambda > 0$ 对于 $d' \in \text{dom } f^H(x; \cdot)$ 有 $d - \lambda d' \in \text{dom } f^c(x; \cdot)$. 由 (5.6.1) 得

$$f^H(x; d) \leq \lambda f^H(x; d') + f^c(x; d - \lambda d'), \quad (5.6.2)$$

因此 $d \in \text{dom } f^H(x; \cdot)$.

(2) 据 (5.6.1) 容易知道, $f^H(x; \cdot)$ 的上图象在 $f^c(x; \cdot)$ 的上图象中稠密, 因此有

$$\liminf_{d' \rightarrow d} f^H(x; d') \leq f^c(x; d).$$

再由定理 5.4.1 知 $f^c(x; d)$ 关于 d 是下半连续的, 故上式为等式.

(3) 已知 $d' \in \text{int dom } f^c(x; \cdot) = \text{dom } f^H(x; \cdot)$, 在 (5.6.2) 中令 $\lambda \rightarrow 0$ 得 $f^c(x; d) = f^H(x; d)$. 由于 $\text{dom } f^H(x; \cdot)$ 是开集, 由定理 5.5.1 可以推得 $f^c: \mathcal{B} \times \mathcal{B} \rightarrow R, (x, d) \mapsto f^c(x; d)$ 在 (x, d') 处是上半连续的.

(4) 由 (3) 和定理 5.4.1 可得.

(5) 根据 (2), 我们有

$$\partial^c f(x) = \{x^* \in \mathcal{B}^* \mid \langle x^*, d \rangle \leq f^H(x; d) \quad \forall d \in \mathcal{B}\}.$$

由已知 $\text{dom } f^H(x; \cdot) = \mathcal{B}$, 只要证明对任意的 $x \in \mathcal{B}$ 有

$$f^H(x; -d) = (-f)^H(x; d).$$

令 $\alpha = f^H(x; -d)$, 对任给定的 $\varepsilon > 0$, 存在 α, β 和 $\eta > 0$, 使得对任意的 $y \in \text{dom } f \cap (x + \alpha B_{\mathcal{B}})$, $t \in (0, \beta)$ 和 $p \in d + \eta B_{\mathcal{B}}$, 有

$$\frac{f(y + tp) - f(y)}{t} \leq \alpha + \varepsilon.$$

取 $\alpha' \in (0, \alpha)$, $\beta' \in (0, \beta)$ 和 $\eta' \in (0, \eta)$, 使得 $\alpha' + \beta'(\|d\|$

$+ \eta') \leq a$, 则对任意的 $x' \in \text{dom} f \cap (x + \alpha' B_{\mathcal{B}})$, $\lambda \in f(x) + \alpha' B_R$, 满足 $\lambda \geq -f(x')$, $t \in (0, \beta')$ 和 $p \in d + \eta' B_{\mathcal{B}}$, 有

$$\frac{-f(x' + tp) - \lambda}{t} \leq \frac{f(x') - f(x' + tp)}{t}.$$

令 $y = x' + tp$, 上式为

$$\frac{-f(x' + tp) - \lambda}{t} \leq \frac{f(y - tp) - f(y)}{t} \leq a + \varepsilon,$$

注意 $y \in \text{dom} f \cap (x + \alpha B)$, 则得 $(-f)^H(x; d) \leq f^H(x; -d)$. 在上式中交换 f 和 $-f$ 即可推出结论. \square

定理 5.6.2 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是实值函数, $x^0 \in \text{dom} f$. 又设 $K \subset \mathcal{B}$ 是闭凸锥, 并且对任意的 $x \in \mathcal{B}$ 和任意的 $y \in K$, 有 $f(x + y) \leq f(x)$.

(1) 对任意的 $d \in K$, 有 $f^c(x^0; d) \leq 0$ 和 $\partial^c f(x^0) \subset K^-$.

(2) 若 $\text{int} K \neq \emptyset$, 则对任意的 $d \in \text{int} K$ 有 $f^H(x^0; d) \leq 0$.

证明 (1) 任给 $\varepsilon > 0$ 和 $\beta > 0$, 对任意的 $x \in x^0 + \varepsilon B$, $\alpha \in f(x^0) + \varepsilon B_R$, 满足 $\alpha \geq f(x)$, $t \in (0, \beta)$ 和 $d \in K$, 我们有

$$\frac{f(x + td) - \alpha}{t} \leq \frac{f(x + td) - f(x)}{t} \leq 0,$$

因此 $f^c(x; d) \leq 0$. 再由定义 5.4.2, 可得 $\partial^c f(x) \subset K^-$.

(2) 因为 $d \in \text{int} K$, 所以存在 $\bar{\delta} > 0$, 使得 $d + \bar{\delta} B \subset \text{int} K$. 据此, 对任意的 $\varepsilon > 0$ 和 $\beta > 0$, $x \in x^0 + \varepsilon B$, $\alpha \geq f(x)$, $t \in (0, \beta)$, 以及 $d' \in d + \bar{\delta} B$, 有

$$\frac{f(x + td') - \alpha}{t} \leq \frac{f(x + td') - f(x)}{t}.$$

根据定理 3.5.1, 即得 $f^H(x_0; d) \leq 0$. \square

定理 5.6.3 设集合 $S \subset \mathcal{B}$ 非空, $x^0 \in S$.

(1) 若 $x^* \in \mathcal{B}^*$ 使得 $\langle x^*, x^0 \rangle = \max_{x \in S} \langle x^*, x \rangle$, 则 $x^* \in$

$N_S^C(x^0)$.

(2) 若 \mathcal{B} 是 Hilbert 空间, $y \notin \text{cl}S$, $x \in \pi_{\text{cl}S}(y)$ (即 $\|x - y\| = \min_{x' \in \text{cl}S} \|y - x'\|$), 则 $y - x \in N_S^c(x)$.

证明 (1) 作函数 $f: \mathcal{B} \rightarrow R$, $f(x) = \langle -x^*, x \rangle$. 因为 f 在 S 上的 x^0 处达到极小值, 根据定理 5.3.2 和推论 4.4.4, 得知 $0 \in \partial^c(f|_S)(x^0) \subset -x^* + N_S^c(x^0)$.

(2) 作函数 $f: \mathcal{B} \rightarrow R$, $f(x) = \|x - y\|$. 由于 f 在 S 上的 x^0 处达到极小值, 我们有

$$\begin{aligned} 0 \in \partial^c(f|_S)(x) &\subset \nabla f(x) + N_S^c(x) \\ &= \frac{x - y}{\|x - y\|} + N_S^c(x), \end{aligned}$$

因此 $y - x \in N_S^c(x)$. \square

定理 5.6.4 设 $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是上半连续的正常函数, $c \in R$, $x^0 \in H(f, c) = \{x \in \mathcal{B} | f(x) \leq c\}$ 使得 $f(x^0) = c$.

(1) $T_{H(f, c)}^D(x^0) \subset \{d \in \mathcal{B} | f^D(x^0; d) \leq 0\}$.

(2) 若存在 $\bar{d} \in \mathcal{B}$ 使得 $f^c(x^0; \bar{d}) < 0$, 则 $\{d \in \mathcal{B} | f^c(x^0; d) \leq 0\} \subset T_{H(f, c)}^D(x^0)$.

证明 (1) 设 $d \in T_{H(f, c)}^D(x^0)$, 则存在序列 $\{d^k\}$, $d^k \rightarrow d$ 和 $\{t_k\}$, $t_k \rightarrow 0^+$ ($k \rightarrow \infty$), 使得 $x^0 + t_k d^k \in H(f, c)$ 对任何 $k = 1, 2, \dots$ 成立. 由此, 我们有 $f(x^0 + t_k d^k) \leq c$ ($k = 1, 2, \dots$), 故

$$\frac{f(x^0 + t_k d^k) - f(x^0)}{t_k} \leq 0,$$

从而 $f^D(x^0; d) \leq 0$.

(2) 设 $a = -f^c(x^0; \bar{d})$, $a > 0$, 则对任意的 $\varepsilon \in (0, a)$, 有 $\alpha > 0$, $\beta > 0$, 使对任意的 $x \in x^0 + \alpha B$ 和 $t \in (0, \beta)$, 存在 $d \in \bar{d} + \varepsilon B$, 使得

$$f(x + td) \leq f(x) + t(-a + \varepsilon).$$

因此, 对任意的 $x \in B_{H(f, c)}(x^0, \alpha)$ 和 $t \in (0, \beta)$, 存在 $\bar{d} \in \bar{d} + \varepsilon B$

使得 $f(x + td) \leq c$, 这即 $x + td \in H(f, c)$, 故 $\bar{d} \in T_{H(f, c)}^c(x^0)$. 设 $d \in \mathcal{B}$ 满足 $f^c(x^0; d) \leq 0$. 对 $\lambda \in (0, 1)$, 令 $d_\lambda = (1 - \lambda)d + \lambda\bar{d}$, 因为 $f^c(x^0; d)$ 关于 d 是凸的, 故 $f^c(x^0; d_\lambda) < 0$. 由此, 我们有 $d_\lambda \in T_{H(f, c)}^c(x^0)$, 令 $\lambda \rightarrow 0$ 得 $d \in T_{H(f, c)}^c(x^0)$. \square

定理 5.6.5 设向量函数 $g: R^n \rightarrow R^m$ 是可微的, $f: R^n \rightarrow R \cup \{+\infty\}$ 是下半连续的正常函数, $x \in g^{-1}(\text{dom} f)$. 若

$$\text{Im} Jg(x) = \text{dom}\{f^c(g(x); \cdot)\} = R^m,$$

则对任意的 $d \in R^n$ 有

$$f^c(g(x); Jg(x)d) \geq (fg)^c(x; d)$$

和

$$\partial^c(fg)(x) \subset Jg(x) \cdot \partial^c f(g(x)).$$

证明 由推论 5.2.8 可推得. \square

定理 5.6.6 设 $f: R^n \rightarrow R \cup \{+\infty\}$ 是下半连续的正常函数, $S \subset R^n$ 是闭集. 若 $x \in S \cap \text{dom} f$, $\text{dom} f^c(x; \cdot) = T_S^c(x) = R^n$, 则对任意的 $d \in T_S^c(x)$, 有 $(f|_S)^c(x; d) \leq f^c(x; d)$ 和

$$\partial^c(f|_S)(x) \subset \partial^c f(x) + N_S^c(x).$$

证明 根据定理 5.2.7 可以推得. \square

设 \mathcal{D} 是 Banach 空间.

定理 5.6.7 设 $g: \mathcal{B} \times \mathcal{D} \rightarrow R \cup \{+\infty\}$ 是正常函数, $f: \mathcal{D} \rightarrow R$,

$$f(y) = \inf\{g(x, y) | x \in \mathcal{B}\}.$$

若 $x_y \in \mathcal{B}$ 使得 $g(x_y, y) = f(y)$, 则对任意的 $p \in \mathcal{D}$ 有

$$f^p(y; d) \leq \inf\{g^p(x_y, y; d, p) | d \in \mathcal{B}\}.$$

证明 设 $d \in \mathcal{B}$, $p \in \mathcal{D}$, 对任给的 $\epsilon > 0$ 和 $\beta > 0$, 当 $d' \in d + \epsilon B_{\mathcal{B}}$, $p' \in p + \epsilon B_{\mathcal{D}}$ 和 $t \in (0, \beta)$ 时, 有

$$\frac{f(y + tp') - f(y)}{t} \leq \frac{g(x_y + td', y + tp') - g(x_y, y)}{t}.$$

据此, 得到

$$f^D(y; p) \leq g^D(x_y, y; d, p), \quad \forall (d, p) \in \mathcal{B} \times \mathcal{D},$$

因而结论成立. \square

注 5.6.1 若 $f(y) = \sup \{g(x, y) | x \in \mathcal{B}\}$, 则有

$$f^D(y; p) \geq \sup \{g^D(x_y, y; d, p) | d \in \mathcal{B}\}.$$

定理 5.6.8 设 $f_i: \mathcal{R}^n \rightarrow R \cap \{+\infty\} (i = 1, \dots, m)$ 是下半连续的 normal 函数, $f(x) = \max \{f_i(x) | i = 1, \dots, m\}$. 又设

$$x^0 \in \bigcap_{i=1}^m \text{int dom } f_i, \quad J(x^0) = \{i = 1, \dots, m | f_i(x^0) = f(x^0)\}.$$

若对任何 $i \in J(x^0)$ 和任何 $d^i \in \mathcal{B}$ 存在 \bar{d} , 使得 $\bar{d} - d^i \in \text{dom } f_i^c(x^0; \cdot)$, 则对任意的 $d \in \mathcal{B}$ 有 $f^c(x^0; d) \leq \max_{i \in J(x^0)} f_i^c(x^0; d)$, 并且

$$\partial^c f(x_0) \subset \overline{\text{co}} \bigcup_{i \in J(x^0)} \partial^c f_i(x^0).$$

证明 已知 $x^0 \in \bigcap_{i=1}^m \text{int dom } f_i$, 当 $i \notin J(x^0)$, 则 $(x^0, f(x^0)) \in \text{int}(\text{epi } f_i)$, 因此切锥 $T_{\text{epi } f_i}^c(x^0, f(x^0)) = \mathcal{B} \times R$. 再由已设, 对所有的 $i = 1, \dots, m$, 存在 \bar{d} 使得 $\bar{d} - d^i \in \text{dom } f_i^c(x^0; \cdot)$, 则对任意的 $(d^i, \lambda_i) \in \mathcal{B} \times R (i = 1, \dots, m)$ 存在 $(d, \bar{\lambda})$, 使得

$$(\bar{d} - d^i, \bar{\lambda} - \lambda_i) \in T_{\text{epi } f_i}^c(x^0, f(x^0)), \quad i = 1, \dots, m,$$

其中 $\bar{\lambda} = \max_{i \in J(x^0)} \{f_i^c(x^0; \bar{d} - d^i) + \lambda_i\}$. 根据推论 5.2.10, 因为

$$\text{epi } f = \bigcap_{i=1}^m \text{epi } f_i, \quad \text{所以有}$$

$$\begin{aligned} & \bigcap_{i \in J(x^0)} T_{\text{epi} f_i}^c(x^0, f(x^0)) \\ &= \bigcap_{i=1}^m T_{\text{epi} f_i}^c(x^0, f(x^0)) \subset T_{\text{epi} f}^c(x^0, f(x^0)), \end{aligned}$$

于是得知结论成立. \square

推论 5.6.9 设 $f_i: R^n \rightarrow R \cup \{+\infty\}$ ($i = 1, \dots, m$) 是下半连续的正常函数. 若 f_i ($i = 1, \dots, m$) 在 $x \in R^n$ 处是 Lipschitz 的, 则对任意的 $d \in R^n$ 有

$$f^c(x; d) \leq \max_{i \in J(x^0)} f_i^c(x; d),$$

$$\partial^c f(x) \subset \overline{\text{co}} \bigcup_{i \in J(x^0)} \partial^c f_i(x).$$

证明 由定理 5.6.8 可推得. \square

定理 5.6.10 设 $f, g: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是正常函数, $x \in \text{dom} f \cap \text{dom} g$. 若 g 在 x 处是正常 H -可切导的, 并且 $d \in \text{dom} f^c(x; \cdot) \cap \text{dom} g^H(x; \cdot) \neq \emptyset$, 则

$$(f + g)^c(x; d) \leq f^c(x; d) + g^c(x; d),$$

$$\partial^c(f + g)(x) \subset \partial^c f(x) + \partial^c g(x).$$

证明 对任意的 $d \in \text{dom} f^c(x; \cdot) \cap \text{dom} g^H(x; \cdot)$, 有

$$(f + g)^c(x; d) \leq f^c(x; d) + g^H(x; d), \quad (5.6.3)$$

设 $d^0 \in \text{dom} f^c(x; \cdot) \cap \text{dom} g^c(x; \cdot)$ 和 $\lambda \in (0, 1)$, 由定理 5.6.1 有 $\text{dom} g^H(x; \cdot) = \text{int dom} f^c(x; \cdot)$, 所以

$$(1 - \lambda)d^0 + \lambda d \in \text{dom} f^c(x; \cdot) \cap \text{dom} g^H(x; \cdot).$$

依据 (5.6.3), 得到

$$\begin{aligned} & (f + g)^c(x; (1 - \lambda)d^0 + \lambda d) \\ & \leq (1 - \lambda)f^c(x; d) + \lambda f^c(x; d) + g^H(x, (1 - \lambda)d^0 + \lambda d). \end{aligned}$$

另外,由(5.6.1)可得

$$g^H(x; (1-\lambda)d^0 + \lambda d) \leq (1-\lambda)g^C(x; d^0) + \lambda g^H(x; d).$$

在以上两式中令 $\lambda \rightarrow 0$, 由(5.6.3)得第1个结论, 再由推论 5.4.4 得第2个结论. \square

最后, 给出有关复合映射的定理如下.

定理 5.6.11 设映射 $\varphi: \mathcal{B} \rightarrow \mathcal{D}$ 是严格可微的, $g: \mathcal{D} \rightarrow R \cup \{+\infty\}$ 是正常函数, $f = g(\varphi(x)); \mathcal{B} \rightarrow R \cup \{+\infty\}$, 并且 g 在 $\varphi(x)$ 处是有限的且具有有限 H -切导数. 若

$$\text{dom} D_S \varphi(x) \cap \text{int} \{d \in \mathcal{B} | g^C(\varphi(x); d) < +\infty\} \neq \emptyset,$$

则

$$\partial^C f(x) \subset D_S \varphi(x) \cdot \partial^C g(\varphi(x)).$$

再若 g 在 x 处是正则的, 则上式成为等式.

证明 设 $y = \varphi(x)$, 作映射 $h: \mathcal{B} \times \mathcal{D} \rightarrow R \cup \{+\infty\}$,

$$h(x', y') = \begin{cases} f(x'), & y' = \varphi(x'); \\ +\infty, & \text{其他.} \end{cases}$$

令 $h = f_1 + f_2$, 其中 f_1 是 φ 图象的指示函数, $f_2(x', y') = g(y')$, $F = D_S \varphi(x)$, $G = \text{graph } \varphi$. 我们证明下面的结论成立:

$$h^C(x, y; d, p) = \begin{cases} f^C(x; d), & p = F(d); \\ +\infty, & \text{其他.} \end{cases}$$

事实上, 有

$$\begin{aligned} & f^C(x, y; d, p) \\ &= \lim_{\substack{(x', y', \alpha) \downarrow (x, y) \\ t \rightarrow 0^+}} \sup \inf_{\substack{d' \rightarrow d \\ p' \rightarrow p}} \frac{h(x' + td', y' + tp') - \alpha}{t}, \end{aligned}$$

其中 $\alpha \geq h(x', y') = f(x')$, $y' = \varphi(x')$, 并且

$$h(x' + td', y' + tp') = \begin{cases} f(x' + td'), & p' = \frac{\varphi(x' + td') - \varphi(x')}{t}; \\ +\infty, & \text{其他.} \end{cases}$$

因为 φ 在 x 处是严格可微的, 我们有

$$\lim_{\substack{x' \rightarrow x, d' \rightarrow d \\ t \rightarrow 0^+}} \sup \frac{\varphi(x' + td') - \varphi(x')}{t} = F(d).$$

因此, 当 $p = F(d)$ 时, 有

$$\lim_{\substack{(x', a) \rightarrow (x, d) \\ t \rightarrow 0^+}} \sup \inf \frac{f(x' + td') - a}{t} = f^c(x; d),$$

并且当 $p \neq F(d)$ 时有 $f^c(x, y; d, p) = +\infty$. 由定义 5.4.2 推知

$$\begin{aligned} \partial^c h(x, y) &= \{(x^*, y^*) \mid h^c(x, y; d, p) \geq \langle x^*, d \rangle \\ &\quad + \langle y^*, p \rangle \quad \forall d \in \mathcal{B}, \forall p \in \mathcal{D}\} \\ &= \{(x^*, y^*) \mid f^c(x, d) \geq \langle x^*, d \rangle \\ &\quad + \langle y^*, F(d) \rangle \quad \forall d \in \mathcal{B}\} \\ &= \{(x^*, y^*) \mid f^c(x, d) \geq \langle x^* + F^*(y^*), d \rangle \quad \forall d \in \mathcal{B}\} \\ &= \{(x^*, y^*) \mid x^* + F^*(y^*) \in \partial^c f(x)\}, \end{aligned}$$

因此得

$$\partial^c f(x) = \{x^* \mid (x^*, 0) \in \partial^c h(x, y)\}. \quad (5.6.4)$$

从 φ 是严格可微的, 我们有 $T_G^c(x, y) = T_G^b(x, y) = \text{graph } \varphi$, 故 f_1 在 (x, y) 处是正则的, 并且有

$$f_1^c(x, y; d, p) = \begin{cases} 0, & p = F(d); \\ +\infty, & \text{其他,} \end{cases}$$

$$\partial^c f_1(x, y) = N_G^c(x, y) = \{(x^*, y^*) \mid x^* = -F^*(y^*)\}. \quad (5.6.5)$$

显然, f_2 在 (x, y) 处是有限正常 H -可切导的, 从而有

$$f_2^c(x, y; d, p) = g^c(y; p),$$

$$\partial^c f_2(x, y) = \{(\mathbf{0}, y^*) \mid y^* \in \partial^c g(y)\}. \quad (5.6.6)$$

由已知得

$$\begin{aligned} & \text{dom} f_1^c(x, y; \cdot, \cdot) \cap \text{dom} f_2^H(x, y; \cdot, \cdot) \\ &= \{(\mathbf{d}, F(\mathbf{d})) \mid F(\mathbf{d}) \in \text{int}\{p \mid g^c(y; p) < +\infty\}\}, \end{aligned}$$

利用定理 5.6.9 有

$$\partial^c h(x, y) \subset \partial^c f_1(x, y) + \partial^c f_2(x, y).$$

根据 (5.6.5)、(5.6.6)、(5.6.4) 和上式即可得到结论. 再若 g 在 x 处是正则的, 从上式不难推出结论对等式情形成立. \square

第 6 章 锥凸集值映射的 锥次微分

从这一章开始,我们研究集值映射的微分问题.集值映射的次微分是现代非光滑分析理论的重要组成部分.

本章考虑在由尖闭凸锥确定偏序的局部凸线性拓扑空间,对锥凸集值映射引进它的锥次微分和锥弱次微分概念.我们指出,本章和下一章介绍的关于集值映射的次微分都是在有效性意义下给出的.可以看到,这些次微分仍然具有一系列重要性质.在本章最后,定义了集值映射的锥方向导数和锥弱方向导数.

§ 6.1 锥次微分和锥弱次微分

设 \mathcal{X} 和 \mathcal{Y} 是局部凸线性拓扑空间, \mathcal{X}^* 和 \mathcal{Y}^* 分别是 \mathcal{X} 和 \mathcal{Y} 的对偶空间, $\langle x^*, x \rangle$ 表示 \mathcal{X} 中连续线性泛函 x^* 在 $x \in \mathcal{X}$ 处的值, $\langle y^*, y \rangle$ 表示连续线性泛函 y^* 在 $y \in \mathcal{Y}$ 处的值.除特别说明外,以下均设 $K \subset \mathcal{Y}$ 是尖闭凸锥, \mathcal{Y} 中的序由 K 确定.设 $x \in \mathcal{X}$, $\phi: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi(x)$ 是集值映射.

先引进集值映射的锥凸性,它是函数凸性的推广.

定义 6.1.1 设 $S \subset \mathcal{X}$ 是非空凸集,集合 $\phi(x) \subset \mathcal{Y}$ ($x \in S$), $\phi: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi(x)$ 是集值映射, $\text{int} K \neq \emptyset$.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} & \lambda\phi(x^1) + (1 - \lambda)\phi(x^2) \\ & \subset \phi(\lambda x^1 + (1 - \lambda)x^2) + K \quad \forall x^1, x^2 \in S, \end{aligned} \quad (6.1.1)$$

则称 ϕ 是 S 上的 K -凸集值映射, 或 ϕ 在 S 上是 K -凸的. 在 (6.1.1) 中令 $x^1 = x^0$ 和 $x^2 = x$, 则称 ϕ 在点 x^0 处(相对于 S)是 K -凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} & \phi(x^1) + (1 - \lambda)\phi(x^2) \\ & \subset \phi(\lambda x^1 + (1 - \lambda)x^2) + \text{int}K \quad \forall x^1, x^2 \in S, \end{aligned} \quad (6.1.2)$$

则称 ϕ 是 S 上的 K -严格凸集值映射, 或 ϕ 在集合 S 上是 K -严格凸的. 在 (6.1.2) 中令 $x^1 = x^0$ 和 $x^2 = x$, 则称 ϕ 在点 x^0 处(相对于 S)是 K -严格凸的.

注 6.1.1 在定义 6.1.1 的(1)中, 当集值映射退化为实值函数 $\phi = f: S \rightarrow R$, 则它在 S 上是 R_+ -凸的, 即 f 是 S 上的凸函数.

下面介绍集合的锥有效性.

定义 6.1.2 设集合 $Y \subset \mathscr{U}$ 非空, $\text{int}K \neq \emptyset$.

(1) 设 $\bar{y} \in Y$, 若不存在 $y \in Y$ 使得

$$\bar{y} - y \in K \setminus \{0\},$$

则称 \bar{y} 是集合 Y 的 K -有效点. Y 的所有 K -有效点组成的集合记作 $\mathscr{E}(Y, K)$.

(2) 设 $\bar{y} \in Y$, 若不存在 $y \in Y$ 使得

$$\bar{y} - y \in \text{int}K,$$

则称 \bar{y} 是集合 Y 的 K -弱有效点. Y 的所有 K -弱有效点组成的集合记作 $\mathscr{E}_w(Y, K)$.

由定义 6.1.2, 显然有

$$\mathscr{E}(Y, K) \subset \mathscr{E}_w(Y, K). \quad (6.1.3)$$

下面给出集值映射的锥次微分和锥弱次微分概念.

定义 6.1.3 设集合 $S \subset \mathscr{X}$ 和 $\phi(x) \subset \mathscr{U} (x \in S)$ 非空. $\phi: S \rightarrow 2^{\mathscr{U}}, x \mapsto \phi(x)$ 是集值映射, 点 $x^0 \in S, y^0 \in \phi(x^0)$, 并且 $x^* \in \mathscr{X}^*,$ 又 $\text{int}K \neq \emptyset$, 给定向量 $p \in \text{int}K$.

(1) 若不存在 $y \in \bigcup_{x \in S} [\psi(x) - \langle x^*, x \rangle p]$ 使得

$$y^0 - \langle x^*, x^0 \rangle p - y \in K \setminus \{0\},$$

则称 x^* 是集值映射 ψ 在点 (x^0, y^0) 处关于向量 p 的 K -次梯度, ψ 在 (x^0, y^0) 处关于 p 的所有 K -次梯度组成的集合称为 ψ 在点 (x^0, y^0) 处关于 p 的 K -次微分, 记作 $\partial\psi(x^0, y^0)_p$. 若 $\partial\psi(x^0, y^0)_p \neq \emptyset$, 则称 ψ 在点 (x^0, y^0) 处关于 p 是 K -次可微的. 若集合

$$\partial\psi(x^0)_p = \{x^* \in \partial\psi(x^0, y^0)_p \mid y^0 \in \psi(x^0)\} \neq \emptyset, \quad (6.1.4)$$

则称集值映射 ψ 在点 x^0 处关于向量 p 是 K -次可微的, 这时称集合 $\partial\psi(x^0)_p$ 是 ψ 在点 x^0 处关于 p 的 K -次微分.

(2) 若不存在 $y \in \bigcup_{x \in S} [\psi(x) - \langle x^*, x \rangle p]$ 使得

$$y^0 - \langle x^*, x^0 \rangle p - y \in \text{int}K,$$

则称 x^* 是集值映射 ψ 在点 (x^0, y^0) 处关于向量 p 的 K -弱次梯度, ψ 在 (x^0, y^0) 处关于 p 的所有 K -弱次梯度组成的集合称为 ψ 在点 (x^0, y^0) 处关于 p 的 K -弱次微分, 记作 $\partial_w\psi(x^0, y^0)_p$. 若 $\partial_w\psi(x^0, y^0)_p \neq \emptyset$, 则称 ψ 在点 (x^0, y^0) 处关于 p 是 K -弱次可微的, 若集合

$$\partial_w\psi(x^0)_p = \{x^* \in \partial_w\psi(x^0, y^0)_p \mid y^0 \in \psi(x^0)\} \neq \emptyset, \quad (6.1.5)$$

则称集值映射 ψ 在点 x^0 处关于向量 p 是 K -弱次可微的, 这时称集合 $\partial_w\psi(x^0)_p$ 是 ψ 在点 x^0 处关于 p 的 K -弱次微分.

按定义 6.1.2 和定义 6.1.3, 则有

$$\begin{aligned} \partial\psi(x^0, y^0)_p &= \left\{ x^* \in \mathcal{X}^* \mid y^0 - \langle x^*, x^0 \rangle p \right. \\ &\quad \left. \in \mathcal{C} \left(\bigcup_{x \in S} [\psi(x) - \langle x^*, x \rangle p], K \right) \right\}, \end{aligned} \quad (6.1.6)$$

$$\begin{aligned} \partial_w \psi(x^0, y^0)_p = & \left\{ x^* \in \mathcal{X}^* \mid y^0 - \langle x^*, x^0 \rangle p \right. \\ & \left. \in \mathcal{E}_w \left(\bigcup_{x \in S} [\psi(x) - \langle x^*, x \rangle p], K \right) \right\}. \end{aligned} \quad (6.1.7)$$

由定义 6.1.3, 显然有

$$\begin{aligned} \partial \psi(x^0, y^0)_p & \subset \partial_w \psi(x^0, y^0)_p, \\ \partial \psi(x^0)_p & \subset \partial_w \psi(x^0)_p. \end{aligned} \quad (6.1.8)$$

注 6.1.2 在定义 6.1.3 的(1)中, 当集值映射退化为实值函数 $\psi = f; S \rightarrow R$, 并且 $K = R_+$ 和 $p = 1$ 时, 有

$$f(x^0) - \langle x^*, x^0 \rangle \in \mathcal{E} \left(\bigcup_{x \in S} [f(x) - \langle x^*, x \rangle], R_+ \right),$$

按定义 6.1.2 的(1), 即不存在 $x \in S$, 使得

$$f(x) - \langle x^*, x \rangle \leq f(x^0) - \langle x^*, x^0 \rangle,$$

或即

$$\langle x^*, x - x^0 \rangle \leq f(x) - f(x^0) \quad \forall x \in S.$$

由此可见, 这时 f 在点 x^0 处关于 $p = 1$ 的 R_+ -次微分, 即是第 3 章中关于凸函数的次微分. 因此, 定义 6.1.3 是凸函数的次梯度和次微分的一种推广.

关于锥次微分和锥弱次微分, 有以下基本结果.

定理 6.1.1 设集合 $S \subset \mathcal{X}$ 非空, $\psi: S \rightarrow 2^Y$ 是集值映射, $x \in S, y \in \psi(x), p \in \text{int} K$.

(1) 若 $x^* \in \partial \psi(x, y)_p$, 则 $y \in \mathcal{E}(\psi(x), K)$; 若 $x^* \in \partial_w \psi(x, y)_p$, 则 $y \in \mathcal{E}_w(\psi(x), K)$.

(2) 若 $\partial \psi(x)_p \neq \emptyset$, 则 $\mathcal{E}(\psi(x), K) \neq \emptyset$; 若 $\partial_w \psi(x)_p \neq \emptyset$, 则 $\mathcal{E}_w(\psi(x), K) \neq \emptyset$.

(3) 若 $y \notin \mathcal{E}(\psi(x), K)$, 则 $\partial \psi(x, y)_p = \emptyset$; 若 $y \notin$

$\mathcal{E}_w(\phi(x), K)$, 则 $\partial_w \phi(x, y)_p = \emptyset$.

证明 (1) 反之, 假设 $y \notin \mathcal{E}(\phi(x), K)$, 由定义 6.1.2 可知存在 $y' \neq y, y' \in \phi(x)$, 使得 $y - y' \in K \setminus \{0\}$, 由此有

$$y - \langle x^*, x \rangle p - (y' - \langle x^*, x \rangle p) \in K \setminus \{0\}.$$

根据定义 6.1.3 的(1)得知 $x^* \notin \partial \phi(x, y)_p$, 导致矛盾. 类似地可证第 2 个结论.

(2) 和 (3) 的第 1 个结论由 (1) 的第 1 个结论可推得, 第 2 个结论由 (1) 的第 2 个结论可推得. \square

定理 6.1.2 设集合 $S \subset \mathcal{X}$ 非空, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S, y \in \phi(x), p \in \text{int}K, \delta > 0$.

(1) 若 $x^* \in \partial \phi(x, y)_p$, 则 $x^* / \delta \in \partial \phi(x, y)_{\delta p}$.

(2) 若 $x^* \in \partial_w \phi(x, y)_p$, 则 $x^* / \delta \in \partial_w \phi(x, y)_{\delta p}$.

证明 由定义 6.1.3 不难推得. \square

定理 6.1.3 设集合 $S \subset \mathcal{X}$ 非空, $x \in S, y \in \phi(x), p \in \text{int}K$.

(1) $0 \in \partial \phi(x, y)_p$ 当且仅当 $y \in \mathcal{E}\left(\bigcup_{x \in S} \phi(x), K\right)$.

(2) $0 \in \partial_w \phi(x, y)_p$ 当且仅当 $y \in \mathcal{E}_w\left(\bigcup_{x \in S} \phi(x), K\right)$.

证明 由定义 6.1.2 容易推得. \square

定理 6.1.4 设集合 $S \subset \mathcal{X}$ 非空, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\phi(x') \neq \emptyset (\forall x' \in S)$. 若 ϕ 在 S 上是 K -严格凸的, $x \in S, y \in \phi(x)$, 则

$$\partial \phi(x, y)_p = \partial_w \phi(x, y)_p, \quad \partial \phi(x)_p = \partial_w \phi(x)_p.$$

证明 记

$$Y(x^*) = \bigcup_{x \in S} [\phi(x) - \langle x^*, x \rangle p], \quad (6.1.9)$$

由 (6.1.3) 我们有

$$\mathcal{E}(Y(x^*), K) \subset \mathcal{E}_w(Y(x^*), K). \quad (6.1.10)$$

下面证明也有

$$\mathcal{E}_w(Y(x^*), K) \subset \mathcal{E}(Y(x^*), K). \quad (6.1.11)$$

事实上, 反之假设有 $y \in \mathcal{E}_w(Y(x^*), K)$ 和 $y \notin \mathcal{E}(Y(x^*), K)$, 从后者按定义 6.1.2 的(1)可知, 存在 $y' \in Y(x^*)$ 使得

$$y - y' \in K \setminus \{\theta\}. \quad (6.1.12)$$

由 $y \in \mathcal{E}_w(Y(x^*), K) \subset Y(x^*)$ 和 $y' \in Y(x^*)$, 根据(6.1.9)知, 存在 $x^1, x^2 \in S$, $y^1 \in \phi(x^1)$, $y^2 \in \phi(x^2)$, 使 $y = y^1 - \langle x^*, x^1 \rangle p$, $y' = y^2 - \langle x^*, x^2 \rangle p$. 因此, 对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} & \lambda y + (1 - \lambda)y' \\ &= \lambda y^1 + (1 - \lambda)y^2 - \langle x^*, \lambda x^1 + (1 - \lambda)x^2 \rangle p, \end{aligned} \quad (6.1.13)$$

由于 ϕ 在 S 上是 K -严格凸的, 按定义 6.1.1 的(2)得知

$$\lambda y^1 + (1 - \lambda)y^2 \in \phi(\lambda x^1 + (1 - \lambda)x^2) + \text{int}K.$$

再由(6.1.13)和(6.1.9)得到

$$\lambda y + (1 - \lambda)y' \in Y(x^*) + \text{int}K.$$

据此, 存在 $y'' \in Y(x^*)$ 和 $v \in \text{int}K$, 使 $y'' + v = \lambda y + (1 - \lambda)y'$, 故有 $y - y'' = (1 - \lambda)(y - y') + v$. 注意到(6.1.12)和 K 是凸锥, 可知

$$y - y'' = (1 - \lambda)(y - y') + v \in K \setminus \{\theta\} + \text{int}K \subset \text{int}K.$$

于是, 由定义 6.1.2 的(2)得到 $y \notin \mathcal{E}_w(Y(x^*), K)$, 这导致与假设矛盾.

由(6.1.10)和(6.1.11)我们有 $\mathcal{E}(Y(x^*), K) = \mathcal{E}_w(Y(x^*), K)$, 由此根据定义 6.1.3 即得结论. \square

§ 6.2 存在性定理

本节研究集值映射的锥次微分和锥弱次微分的存在性.

设集合 $S \subset \mathcal{X}$, $\phi: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi(x)$ 是集值映射, $K \subset \mathcal{Y}$ 是内部非空的尖闭凸锥.

定义 6.2.1 设集合 $S \subset \mathcal{X}$ 非空, 点 $x^0 \in S$, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射. 若存在点 $a \in \mathcal{Y}$ 和集合 $T \subset \mathcal{Y}$, 使得 $\phi(x^0) \subset T \subset a - K$ 和 $\phi^{-1}(T) = \{x \in S | \phi(x) \subset T\}$ 是开集, 则称 ϕ 在点 x^0 处是 K -有界逆开的.

记集值映射 ϕ 在 S 上的 K -上图象为

$$K\text{-epi}\phi = \{(x, y) \in \mathcal{X} \times \mathcal{Y} | x \in S, y \in \phi(x) + K\}. \quad (6.2.1)$$

引理 6.2.1 设集合 $S \subset \mathcal{X}$ 非空, $x \in S$. 若集值映射 $\phi: S \rightarrow 2^{\mathcal{Y}}$ 在 x 处是 K -有界逆开的, 则 $\text{int}(K\text{-epi}\phi) \neq \emptyset$.

证明 因 ϕ 在 x 处是 K -有界逆开的, 按定义 6.2.1 可知, 存在 $a \in \mathcal{Y}$ 和 $T \subset \mathcal{Y}$, 使得对任意的 $x \in \phi^{-1}(T)$, 取 $y_x \in \phi(x)$, 则存在 $p_x \in K$ 使 $y_x = a - p_x$. 设 $\bar{y} = a + p$, 则 $\bar{y} - a = p \in \text{int}K$, 于是存在 θ 点的邻域 $U(\theta) \subset \mathcal{Y}$ 有

$$y + \bar{y} - a \in K \quad \forall y \in U(\theta).$$

因为 K 是凸锥, 所以从上式可得

$$y + \bar{y} - y_x = y + \bar{y} - a + p_x \subset K$$

$$\forall x \in \phi^{-1}(T), \forall y \in U(\theta),$$

或即

$$y + \bar{y} \subset y_x + K \subset \phi(x) + K$$

$$\forall x \in \phi^{-1}(T), \forall y \in U(\theta).$$

据此, 由 (6.2.1) 得

$$(x, y) \in K\text{-epi}\phi \quad \forall x \in \phi^{-1}(T), \forall y \in U(\theta) + \bar{y}.$$

由于 $\phi^{-1}(T)$ 和 $U(\theta) + \bar{y}$ 是非空开集, 因此 $\text{int}(K\text{-epi}\phi) \neq \emptyset$. \square

引理 6.2.2 设 $S \subset \mathcal{X}$ 是开集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射. 若对任意的 $x \in S$, 存在 $a \in \mathcal{Y}$, 使得 $\psi(x) \subset a - K$, 则 $\text{int}(K\text{-epi}\psi) \neq \emptyset$.

证明 设 $T = a - K$, 则 ψ 在 x 处是 K -有界逆开的. 由引理 6.2.1, 即得 $\text{int}(K\text{-epi}\psi) \neq \emptyset$. \square

定义 6.2.2 设集合 $S \subset \mathcal{X}$ 非空, 点 $x^0 \in S$, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射. 若存在映射 $\varphi: S \rightarrow \mathcal{Y}$, 使得 $\varphi(x) \in \psi(x) (\forall x \in S)$, 并且 φ 在点 x^0 的邻域 $U(x^0)$ 内是连续的, 则称 ψ 在点 x^0 处是连通的.

引理 6.2.3 设集合 $S \subset \mathcal{X}$ 非空, $x \in S$, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射. 若 ψ 在 x 处是连通的, 则 $\text{int}(K\text{-epi}\psi) \neq \emptyset$.

证明 因为 φ 在点 x 的邻域 $U(x)$ 内是连续的, 所以 $\varphi(U(x))$ 是开集. 由此, 我们有 $\{U(x), \varphi(U(x))\} \subset K\text{-epi}\psi$, 故 $\text{int}(K\text{-epi}\psi) \neq \emptyset$. \square

以下给出锥凸集值映射的锥次微分的存在性定理.

定理 6.2.4 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\psi(x') \neq \emptyset (\forall x' \in S)$, 并且 $x \in \text{int}S$, $p \in \text{int}K$. 又设 ψ 在 S 上是 K -凸的, 在 x 处是 K -严格凸的和 K -有界逆开的, 并且 $\mathcal{S}(\psi(x), K) \neq \emptyset$.

(1) 若 $y \in \mathcal{S}(\psi(x), K)$, 则 $\partial\psi(x, y)_p \neq \emptyset$.

(2) $\partial\psi(x)_p \neq \emptyset$.

证明 (1) 首先证明 $K\text{-epi}\psi$ 是 $\mathcal{X} \times \mathcal{Y}$ 中的凸集. 设取任意的 $(x^1, y^1), (x^2, y^2) \in K\text{-epi}\psi$, 即 $y^1 \in \psi(x^1) + K, y^2 \in \psi(x^2) + K$. 因为 S 是凸的, 设 $\lambda \in (0, 1)$, 则 $\lambda x^1 + (1 - \lambda)x^2 \in S$. 又因为 K 也是凸的, ψ 在 S 上是 K -凸的, 依据定义 6.1.1 的(1)有

$$\begin{aligned} \lambda y^1 + (1 - \lambda)y^2 &\in \lambda\psi(x^1) + (1 - \lambda)\psi(x^2) + K \\ &\subset \psi(\lambda x^1 + (1 - \lambda)x^2) + K, \end{aligned}$$

从而 $(\lambda x^1 + (1 - \lambda)x^2, \lambda y^1 + (1 - \lambda)y^2) \in K\text{-epi}\psi$, 故 $K\text{-epi}\psi$ 是凸集. 由 ψ 在 x 处是 K -有界逆开的, 根据引理 6.2.1 得知

$$\text{int}(K\text{-epi}\psi) \neq \emptyset. \quad (6.2.2)$$

已知 $y \in \mathcal{E}(\psi(x), K)$, 我们证明

$$(x, y) \notin \text{int}(K\text{-epi}\psi). \quad (6.2.3)$$

事实上, 若 $(x, y) \in \text{int}(K\text{-epi}\psi)$, 则存在点 θ 的邻域 $U_1(\theta) \subset \mathcal{W}$ 使得

$$(x, y' + y) \in K\text{-epi}\psi \quad \forall y' \in U_1(\theta). \quad (6.2.4)$$

因为 $p \neq \theta$ 和 $U_1(\theta)$ 是吸收集, 所以存在 $\mu > 0$, 使得 $-\mu p \in U_1(\theta)$. 由此, 从 (6.2.4) 得 $y' - \mu p \in \psi(x) + K$, 于是对应存在 $\bar{y} \in \psi(x)$, $\bar{p} \in K$ 使得 $\bar{y} + \bar{p} = y - \mu p$, 即

$$y - \bar{y} = \bar{p} + \mu p \in \text{int}K \subset K \setminus \{\theta\}.$$

根据定义 6.1.2 的(1), 即知 $y \notin \mathcal{E}(\psi(x), K)$. 这导致与定理的条件相矛盾.

现在由 (6.2.2) 和 (6.2.3), 根据定理 1.3.7 的(1), 可知存在 $(x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^*$, $(x^*, y^*) \neq \theta$, 使得

$$\begin{aligned} \langle x^*, x \rangle + \langle y^*, y \rangle &\geq \langle x^*, x' \rangle + \langle y^*, y' \rangle \\ \forall x' \in S, \forall y' \in \psi(x') + K. \end{aligned} \quad (6.2.5)$$

在 (6.2.5) 中 $y^* \neq \theta$, 因为否则, 若 $y^* = \theta$, 由 $(x^*, y^*) \neq \theta$ 知 $x^* \neq \theta$. 代入 (6.2.5) 得到

$$\langle x^*, x - x' \rangle \geq 0 \quad \forall x' \in S. \quad (6.2.6)$$

于是, 存在 $\bar{x} \in \mathcal{X}$ 使得 $\mu = \langle x^*, \bar{x} \rangle > 0$. 再由 $x \in \text{int}S$ 知存在 x 的邻域 $U_2(x) \subset \text{int}S$, 因此有充分大的 $M > 0$ 使 $x + (\mu/M)\bar{x} \in U_2(x)$. 将它代入 (6.2.6), 有 $\langle x^*, -(\mu/M)\bar{x} \rangle \geq 0$, 这导致 $-\mu^2 \geq 0$, 产生矛盾. 现任取 $q \in K$, 因为 $y \in \psi(x)$, 故 $y + q \in \psi(x) + K$, 代入 (6.2.5) 有

$$\langle x^*, x \rangle + \langle y^*, y \rangle \geq \langle x^*, x \rangle + \langle y^*, y + q \rangle,$$

从而得

$$\langle -y^*, q \rangle \geq 0 \quad \forall q \in K.$$

注意到 $y^* \neq 0$, 有 $-y^* \in K^* \setminus \{0\}$. 因为 $p \in \text{int}K$, 由定理 1.4.11 的(1)得 $\langle -y^*, p \rangle > 0$. 令

$$\delta = \langle -y^*, p \rangle, \quad (6.2.7)$$

进一步证明下式成立:

$$\langle x^*, x \rangle + \langle y^*, y \rangle > \langle x^*, x' \rangle + \langle y^*, y' \rangle, \quad (6.2.8)$$

其中 $(x, y) \neq (x', y') (\forall x' \in S, \forall y' \in \phi(x') + K)$. 事实上, 假若(6.2.8)不成立, 从(6.2.5)可知存在 $(\bar{x}, \bar{y} + \bar{q}) \in K\text{-epi}\phi$, $\bar{y} \in \phi(\bar{x})$ 和 $(\bar{x}, \bar{y} + \bar{q}) \neq (x, y)$, 使得

$$\langle x^*, x \rangle + \langle y^*, y \rangle = \langle x^*, \bar{x} \rangle + \langle y^*, \bar{y} \rangle + \bar{q}. \quad (6.2.9)$$

对任意的 $\lambda \in (0, 1)$ 令

$$x_\lambda = \lambda \bar{x} + (1 - \lambda)x, \quad (6.2.10)$$

$$y_\lambda = \lambda \bar{y} + (1 - \lambda)y, \quad (6.2.11)$$

因为 $\bar{y} \in \phi(\bar{x})$, $y \in \phi(x)$, ϕ 在 x 处是 K -严格凸的, 我们有

$$y_\lambda = \lambda \bar{y} + (1 - \lambda)y \in \phi(\lambda \bar{x} + (1 - \lambda)x) + \text{int}K.$$

又因为 S 是凸集, 故 $x_\lambda \in S$. 设 $y_\lambda = \eta + \gamma$, 其中 $\eta \in \phi(x_\lambda)$, $\gamma \in \text{int}K$. 代入(6.2.5)得

$$\langle x^*, x \rangle + \langle y^*, y \rangle \geq \langle x^*, x_\lambda \rangle + \langle y^*, \eta \rangle. \quad (6.2.12)$$

由(6.2.5), (6.2.10)和(6.2.11), 推知有

$$\begin{aligned} & \langle x^*, x_\lambda \rangle + \langle y^*, y_\lambda \rangle \\ &= \langle x^*, \lambda \bar{x} + (1 - \lambda)x \rangle + \langle y^*, \lambda \bar{y} + (1 - \lambda)y \rangle \\ &= \lambda [\langle x^*, \bar{x} \rangle + \langle y^*, \bar{y} \rangle] \end{aligned}$$

$$\begin{aligned}
 & + (1 - \lambda) [\langle x^*, x \rangle + \langle y^*, y \rangle] \\
 & = \langle x^*, x \rangle + \langle y^*, y \rangle - \lambda \langle y^*, \bar{q} \rangle.
 \end{aligned}$$

据此,由(6.2.12)得知

$$\langle y^*, y_\lambda \rangle + \lambda \langle y^*, \bar{q} \rangle \geq \langle y^*, \eta \rangle,$$

再由 $\langle -y^*, \bar{q} \rangle \geq 0$, 得到

$$\langle y^*, y_\lambda \rangle \geq \langle y^*, \eta \rangle. \quad (6.2.13)$$

据定理 1.4.11 的(1)知 $\langle y^*, \gamma \rangle < 0$, 从而有

$$\langle y^*, y_\lambda \rangle = \langle y^*, \eta + \gamma \rangle < \langle y^*, \eta \rangle,$$

它与(6.2.13)相矛盾.

最后,假设若 $\partial\phi(x, y)_* = \emptyset$, 由定理 6.1.3 知, 对于 $\delta > 0$ 有 $\partial\phi(x, y)_\delta = \emptyset$, 即 $x^* \notin \partial\phi(x, y)_\delta$. 根据定义 6.1.3 的(1), 存在 $\bar{x} \in S$, $\bar{y} \in \phi(\bar{x})$, $(\bar{x}, \bar{y}) \neq (x, y)$, 使得

$$y - \langle x^*, x \rangle \frac{p}{\delta} - \langle \bar{y} - \langle x^*, \bar{x} \rangle \frac{p}{\delta} \in K \setminus \{0\}.$$

由 $-y^* \in K^* \setminus \{0\}$ 即可推得

$$\langle x^*, x \rangle + \langle y^*, y \rangle \leq \langle x^*, \bar{x} \rangle + \langle y^*, \bar{y} \rangle,$$

这与(6.2.6)相矛盾.

(2) 由已证明的(1)依据定义 6.1.3 的(1), 得知 $\partial\phi(x)_* \neq \emptyset$.]

设 \mathcal{V} 是向量空间.

定义 6.2.3 设集合 $S \subset \mathcal{V}$ 非空, $K \subset \mathcal{V}$ 是非平凡锥. 若 S 的每一个形如 $\{(y - \text{cl}K)^c \mid y \in S\}_{c \in \Gamma}$ 的开覆盖都有一个有限子覆盖 (Γ 是指标集), 则称 S 是 K -半紧的.

引理 6.2.5 设集合 $Y \subset \mathcal{V}$ 非空, $K \subset \mathcal{V}$ 是尖闭凸锥. 若 Y 是 K -半紧的, 则 $\mathcal{E}(Y, K) \neq \emptyset$.

证明 见文献[27]的定理 3.3.1.

定理 6.2.6 设 $S \subset \mathcal{X}$ 是非空开凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\phi(x') \neq \emptyset$ ($\forall x' \in S$), $x \in S, p \in \text{int}K$. 又设 ϕ 在 S 上是 K -凸的, 在 x 处是 K -严格凸的, 并且存在 $a \in \mathcal{Y}$ 使得对任意的 $x' \in S$, 有 $\phi(x') \subset a - K$, 以及 $\phi(x)$ 是 K -半紧的.

(1) 若 $y \in \mathcal{S}(\phi(x), K)$, 则 $\partial \phi(x, y)_p \neq \emptyset$.

(2) $\partial \phi(x)_p \neq \emptyset$.

证明 由引理 6.2.2 和引理 6.2.5, 以及定理 6.2.4 可证得. \square

定理 6.2.7 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\phi(x') \neq \emptyset$ ($\forall x' \in S$), $x \in \text{int}S, p \in \text{int}K$. 又设 ϕ 在 S 上是 K -凸的, 在 x 处是 K -严格凸的和连通的, 并且 $\phi(x)$ 是 K -半紧的.

(1) 若 $y \in \mathcal{S}(\phi(x), K)$, 则 $\partial \phi(x, y)_p \neq \emptyset$.

(2) $\partial \phi(x)_p \neq \emptyset$.

证明 由引理 6.2.3, 引理 6.2.5 和定理 6.2.4, 可以推得. \square

以下给出锥凸集值映射的锥弱次微分的存在性定理.

定理 6.2.8 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\phi(x') \neq \emptyset$ ($\forall x' \in S$), $x \in \text{int}S, p \in \text{int}K$. 又设 ϕ 在 S 上是 K -凸的, 在 x 处是 K -有界逆开的, 并且 $\mathcal{S}_w(\phi(x), K) \neq \emptyset$.

(1) 若 $y \in \mathcal{S}_w(\phi(x), K)$, 则 $\partial_w \phi(x, y)_p \neq \emptyset$.

(2) $\partial_w \phi(x)_p \neq \emptyset$.

证明 (1) 类似于定理 6.2.4, 可以证明 $K\text{-epi}\phi$ 是内部非空的凸集, 并且对于 $y \in \mathcal{S}_w(\phi(x), K)$ 有 $(x, y) \notin \text{int}(K\text{-epi}\phi)$. 据此, 由定理 1.3.7 的(1)可知, 存在 $(x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^*, (x^*, y^*) \neq \theta$, 使得

$$\langle x^*, x \rangle + \langle y^*, y \rangle \geq \langle x^*, x' \rangle + \langle y^*, y' \rangle$$

$$\forall x' \in S, \forall y' \in \phi(x') + K. \quad (6.2.14)$$

同定理 6.2.4 的证明, 还可得 $y^* \neq \theta, -y^* \in K^* \setminus \{\theta\}$. 由于 $p \in \text{int}K$, 根据定理 1.4.11 的(1)有

$$\delta = \langle y^*, p \rangle > 0. \quad (6.2.15)$$

现在用反证法. 假设 $\partial_w \phi(x, y)_p = \emptyset$, 则由定理 6.1.3 的(2) 得知 $\partial_w \phi(x, y)_{\frac{p}{\delta}} = \emptyset$, 于是由定义 6.1.3 的(2) 可知存在 $\bar{x} \in S, \bar{y} \in \phi(\bar{x})$, 使得

$$y - \langle x^*, x \rangle \frac{p}{\delta} - \left(\bar{y} - \langle x^*, \bar{x} \rangle \frac{p}{\delta} \right) \in \text{int}K.$$

再根据定理 1.4.11, 我们有

$$\left\langle -y^*, y - \bar{y} - \langle x^*, x - \bar{x} \rangle \frac{p}{\delta} \right\rangle > 0,$$

利用(6.2.15)整理上式, 得

$$\langle x^*, x \rangle + \langle y^*, y \rangle < \langle x^*, \bar{x} \rangle + \langle y^*, \bar{y} \rangle,$$

这导致与(6.2.14)相矛盾.

(2) 由(1)依据定义 6.1.3 的(2), 即得 $\partial_w \phi(x)_p \neq \emptyset$. \square

定理 6.2.9 设 $S \subset \mathcal{X}$ 是非空开凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\phi(x') \neq \emptyset$ ($\forall x' \in S$), $x \in S, p \in \text{int}K$. 又设 ϕ 在 S 上是 K -凸的, 存在 $a \in \mathcal{Y}$ 使得对任意的 $x' \in S$ 有 $\phi(x') \subset a - K$, 并且 $\phi(x')$ 是 K -半紧的.

(1) 若 $y \in \mathcal{S}_w(\phi(x), K)$, 则 $\partial_w \phi(x, y)_p \neq \emptyset$.

(2) $\partial_w \phi(x)_p \neq \emptyset$.

证明 由引理 6.2.2 可知 $\text{int}(K\text{-epi}\phi) \neq \emptyset$, 又由引理 6.2.5 有 $\mathcal{S}_w(\phi(x), K) \neq \emptyset$. 于是, 由定理 6.2.8 即得结论. \square

定理 6.2.10 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $\phi(x') \neq \emptyset$ ($\forall x' \in S$), $x \in \text{int}S, p \in \text{int}K$. 又设 ϕ 在 S 上是 K -凸的, 在 x 处是连通的, 并且 $\phi(x)$ 是 K -半紧的.

(1) 若 $y \in \mathcal{S}_w(\phi(x), K)$, 则 $\partial_w \phi(x, y)_p \neq \emptyset$.

(2) $\partial_w \phi(x)_p \neq \emptyset$.

证明 由引理 6.2.3、引理 6.2.5 和定理 6.2.8 可得证. \square

下面给出锥次微分和锥弱次微分存在的充分条件.

定理 6.2.11 设集合 $S \subset \mathscr{X}$ 非空, $x \in S$, $\phi: S \rightarrow 2^*$ 是集值映射, $y \in \phi(x)$. 又 $K^* \subset \mathscr{X}^*$ 是 K 的对偶锥, $x^* \in \mathscr{X}^*$, $p \in \text{int}K$.

(1) 若存在 $q^* \in \text{int}K^*$ 使得

$$\begin{aligned} \langle q^*, y' - y \rangle &\geq \langle x^*, x' - x \rangle \langle q^*, p \rangle \\ \forall x' \in S, \forall y' \in \phi(x'), \end{aligned} \quad (6.2.16)$$

则 $x^* \in \partial\phi(x, y)_p$.

(2) 若存在 $q^* \in K^* \setminus \{0\}$ 使得

$$\begin{aligned} \langle q^*, y' - y \rangle &\geq \langle x^*, x' - x \rangle \langle q^*, p \rangle \\ \forall x' \in S, \forall y' \in \phi(x'), \end{aligned} \quad (6.2.17)$$

则 $x^* \in \partial_\omega\phi(x, y)_p$.

证明 (1) 反之, 假设 $x^* \notin \partial\phi(x, y)_p$, 由定义 6.1.3 的(1)知存在 $\bar{x} \in S$, $\bar{y} \in \phi(\bar{x})$ 使得

$$y - \langle x^*, x \rangle p - (\bar{y} - \langle x^*, \bar{x} \rangle p) \in K \setminus \{0\}.$$

根据定理 1.4.11 的(2)有

$$\langle q^*, y - \langle x^*, x \rangle p - (\bar{y} - \langle x^*, \bar{x} \rangle p) \rangle > 0,$$

整理上式得

$$\langle q^*, y - \bar{y} \rangle > \langle x^*, x - \bar{x} \rangle \langle q^*, p \rangle,$$

这导致与(6.2.16)矛盾.

(2) 用反证法. 利用定理 1.4.11, 类似于(1)的证明可得证. \square

下面的定理说明, 在凸性条件下, (6.2.17)还是锥次微分和锥弱次微分存在的必要条件.

定理 6.2.12 设集合 $S \subset \mathscr{X}$ 非空, $x \in S$, $\phi: S \rightarrow 2^*$ 是集值映射, $y \in \phi(x)$, $x^* \in \mathscr{X}^*$, $p \in \text{int}K$, K^* 是 K 的对偶锥. 若 ϕ 在 S 上是 K -凸的, 并且 $x^* \in \partial_\omega\phi(x, y)_p$, 则存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' - y \rangle \geq \langle x^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y' \in \phi(x').$$

证明 作集合

$$E = \{y' \in \mathcal{Y} \mid y' = y - \bar{y} + \langle x^*, \bar{x} - x \rangle p - \bar{q},$$

$$\bar{x} \in S, \bar{y} \in \phi(\bar{x}), \bar{q} \in K\}, \quad (6.2.18)$$

则由 S 和 K 是凸的, 以及 ϕ 是 K -凸的可以推知 E 是凸集. 下面证明

$$\text{int}K \cap E = \emptyset. \quad (6.2.19)$$

事实上, 若存在 $q' \in \text{int}K$ 使 $q' \in E$, 由 (6.2.18) 则有

$$q' = y - \bar{y} + \langle x^*, \bar{x} - x \rangle p - \bar{q},$$

即

$$y - \langle x^*, x \rangle p - (\bar{y} - \langle x^*, \bar{x} \rangle p) = \bar{q} + q' \in \text{int}K.$$

于是, 按定义 6.1.3 的 (2) 得知 $x^* \notin \partial_w \phi(x, y)_p$, 它与所设矛盾. 现在由 (6.2.19) 根据定理 1.3.7 可知, 存在 $q^* \in \mathcal{Y}^* \setminus \{\theta\}$, 使得

$$\langle q^*, q \rangle \geq \langle q^*, y' \rangle \quad \forall q \in K, \forall y' \in E.$$

由 $\theta \in K \cap E$, 从上式得知对任意的 $q \in K$ 有 $\langle q^*, q \rangle \geq 0$, 故 $q^* \in K^* \setminus \{\theta\}$. 再由 $0 \geq \langle q^*, y' \rangle (\forall y' \in E)$, 即得 (6.2.17). \square

特别地, 对任意的 $q^* \in K^* \setminus \{\theta\}$, 作集值函数 $\theta: S \rightarrow 2^R, x \mapsto \theta(x), \theta(x) = \langle q^*, \phi(x) \rangle = \{\langle q^*, y \rangle \mid y \in \phi(x)\} (\forall x \in S)$. 取 R 中的尖闭凸锥为 R_+ , 显然因 $R_+ \setminus \{\theta\} = \text{int}R_+$, 再令 $p = 1$, 则由定义 6.1.3 可知, 对于 $x \in S, y \in \phi(x)$, 我们有

$$\partial\theta(x, \langle q^*, y \rangle)_1 = \partial_w\theta(x, \langle q^*, y \rangle)_1.$$

定理 6.2.13 设集合 $S \subset \mathcal{X}$ 非空, $x \in S, \phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $y \in \phi(x)$. 又设 K^* 是 K 的对偶锥, ϕ 在 S 上是 K -凸的, $p \in \text{int}K$.

$$(1) \bigcup_{q^* \in \text{int} K^*} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1 \subset \partial \phi(x, y)_p.$$

$$(2) \bigcup_{q^* \in K^* \setminus \{0\}} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1 = \partial_w \phi(x, y)_p.$$

证明 (1) 对任意的 $q^* \in \text{int} K^*$, 设

$$x^* \in (\langle q^*, p \rangle)^{-1} \partial \theta(\bar{x}, \langle q^*, y \rangle)_1,$$

则

$$\langle q^*, p \rangle x^* \in \partial \theta(x, \langle q^*, y \rangle)_1.$$

由 ϕ 在 S 上是 K -凸的, 容易证明 θ 在 S 上是 R_+ -凸的. 于是, 由定理 6.2.12 知存在 $\lambda \in R_+ \setminus \{0\}$, 使得

$$\lambda(\langle q^*, y' - \bar{y} \rangle) \leq \langle \langle q^*, p \rangle x^*, x' - x \rangle \lambda$$

$$\forall x' \in S, \forall y' \in \theta(x'),$$

即

$$\langle q^*, y' - \bar{y} \rangle \geq \langle x^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y' \in \phi(x').$$

由此, 根据定理 6.2.11 得到 $x^* \in \partial \phi(x, y)_p$.

(2) 类似于(1)中的证明, 由定理 6.2.11 和定理 6.2.12 可以推得

$$\bigcup_{q^* \in K^* \setminus \{0\}} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1 \subset \partial_w \phi(x, y)_p.$$

其次, 设 $x^* \in \partial_w \phi(x, y)_p$, 由定理 6.2.12 知存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' - \bar{y} \rangle \geq \langle x^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y' \in \phi(x').$$

于是, 根据定理 6.2.11 得 $x^* \langle q^*, p \rangle \in \partial \theta(x, \langle q^*, y \rangle)_1$, 即有 $x^* \in (\langle q^*, p \rangle)^{-1} \partial \theta(x, \langle q^*, y \rangle)_1$. \square

设 $f: S \rightarrow R$ 是实值函数, $f(x) = \inf \langle q^*, \phi(x) \rangle$, 其中 $q^* \in K^* \setminus \{0\}$.

定理 6.2.14 设 $S \subset \mathcal{X}$ 是非空凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S$, $y \in \phi(x)$, 并且 $p \in \text{int}K$, K^* 是 K 的对偶锥. 若 ϕ 在 S 上是 K -凸的, $f(x) = \langle q^*, y \rangle (\forall q^* \in K^* \setminus \{0\})$, 对任意的 $x' \in S$, $\phi(x')$ 是非空紧集, 则对任意的 $q^* \in K^* \setminus \{0\}$, 函数 f 在 S 上是凸的, 并且

$$(1) \bigcup_{q^* \in \text{int}K^*} \frac{1}{\langle q^*, p \rangle} \partial f(x) \subset \partial \phi(x, y)_p,$$

$$(2) \bigcup_{q^* \in K^* \setminus \{0\}} \frac{1}{\langle q^*, p \rangle} \partial f(x) = \partial_w \phi(x, y)_p,$$

其中 $\partial f(x) = \partial f(x, \langle q^*, y \rangle)_1$.

证明 任取 $x^1, x^2 \in S$, $\lambda \in (0, 1)$, 因为已知对任何连续线性泛函 $q^* \in K^* \setminus \{0\}$, $\phi(x) (\forall x \in S)$ 是紧集, 所以存在 $y^1 \in \phi(x^1)$, $y^2 \in \phi(x^2)$, 使得 $f(x^1) = \langle q^*, y^1 \rangle$, $f(x^2) = \langle q^*, y^2 \rangle$. 由 S 是凸的, ϕ 在 S 上是 K -凸的, 则有

$$\lambda y^1 + (1 - \lambda)y^2 \in \phi(\lambda x^1 + (1 - \lambda)x^2) + K.$$

因此, 存在 $y_\lambda \in \phi(\lambda x^1 + (1 - \lambda)x^2)$ 和 $q \in K$, 使得

$$\lambda y^1 + (1 - \lambda)y^2 = y_\lambda + q.$$

据此, 由 $q^* \in K^* \setminus \{0\}$ 得知

$$\langle q^*, \lambda y^1 \rangle + \langle q^*, (1 - \lambda)y^2 \rangle = \langle q^*, y_\lambda \rangle + \langle q^*, q \rangle,$$

即

$$\begin{aligned} \lambda f(x^1) + (1 - \lambda)f(x^2) &\geq \langle q^*, y_\lambda \rangle \\ &\geq f(\lambda x^1 + (1 - \lambda)x^2), \end{aligned}$$

故 f 在 S 上是凸的. 由定理 6.2.11 和定理 6.2.12 有

$$\partial f(x) = \{x^* \mid \langle x^*, x' - x \rangle \leq f(x') - f(x) \forall x' \in S\}. \quad (6.2.20)$$

现在证明 $\partial f(x) = \partial \theta(x, \langle q^*, y \rangle)_1$. 设 $x^* \in \partial f(x)$, 由 (6.2.20) 和 f 的定义可知, 对任意的 $x' \in S$ 和任意的 $y' \in \phi(x')$ 有

$$\langle q^*, y' \rangle - \langle q^*, y \rangle \geq f(x') - f(x) \geq \langle x^*, x' \rangle.$$

由此, 根据定理 6.2.11 得 $x^* \in \partial \theta(x, \langle q^*, y \rangle)_1$. 反之, 设 $x^* \in \partial \theta(x, \langle q^*, y \rangle)_1$, 由定理 6.2.12 知对任意的 $x' \in S$ 和任意的 $y' \in \phi(x')$ 有

$$\langle q^*, y' \rangle - \langle q^*, y \rangle \geq \langle x^*, x' - x \rangle.$$

注意到 $\phi(x')$ 是紧集, 则存在 y_x 使得 $f(x') = \langle q^*, y_x \rangle$, $y_x \in \phi(x')$, 故由上式得知

$$\langle x^*, x' - x \rangle \leq f(x') - f(x) \quad \forall x' \in S,$$

再由 (6.2.20) 便得 $x^* \in \partial f(x)$. 根据以上结果, 由定理 6.2.13, 即得 (1) 和 (2) 的结论. \square

对于 Banach 空间的情况, 有以下一些结果.

定理 6.2.15 设 \mathcal{B} 是 Banach 空间, $S \subset \mathcal{B}$ 是非空凸集, $\phi: S \rightarrow 2^\infty$ 是集值映射, $\phi(x') \neq \emptyset$ ($\forall x' \in S$). 又设 $x \in S$, $y \in \phi(x)$, $p \in \text{int} K$, K^* 是 K 的对偶锥. 若 ϕ 在 S 上是 K -凸的, 则

$$(1) \bigcup_{q^* \in K_1^*} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1 \subset \partial \phi(x, y)_p,$$

$$(2) \bigcup_{q^* \in K_2^*} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1 = \partial_w \phi(x, y)_p,$$

其中 $K_1^* = \{q^* \in \text{int} K^* \mid \|q^*\|_* = 1\}$,

$$K_2^* = \{q^* \in K^* \setminus \{0\} \mid \|q^*\|_* = 1\},$$

$\|\cdot\|_*$ 是对偶空间 \mathcal{B}^* 的范数.

证明 (1) 由定理 6.2.13 结论显然成立.

(2) 据定理 6.2.13 的 (2), 只要证明

$$\partial_w \phi(x, y)_p \subset \bigcup_{q^* \in K_2^*} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1.$$

设 $x^* \in \partial_{\omega} \psi(x, y)_p$, 由定理 6.2.12, 存在 $q^* \in K^* \setminus \{0\}$ 使得

$$\langle q^*, y' - y \rangle \geq \langle x^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

因为 $\|q^*\|_* \neq 0$, 在上式两端同除一个 $\|q^*\|_*$, 令 $q_1^* = q^* / \|q^*\|_*$, 得

$$\langle q_1^*, y' - y \rangle \geq \langle x^*, x' - x \rangle \langle q_1^*, p \rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

因此, 得知 $q_1^* \in K_1^*$, $x^* \langle q_1^*, p \rangle \in \partial f(x, \langle q_1^*, y \rangle)_1$. 定理得证. \square

定理 6.2.16 设 \mathcal{B} 是 Banach 空间, $S \subset \mathcal{B}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{B}}$ 是集值映射. 又设 $x \in S$, $y \in \psi(x)$, $p \in \text{int}K$, K^* 是 K 的对偶锥. 若 ψ 在 S 上是 K -凸的, $f(x) = \langle q^*, y \rangle (\forall q^* \in K^* \setminus \{0\})$, 对任意的 $x' \in S$, $\psi(x')$ 是紧集, 则

$$(1) \bigcup_{q^* \in K_1^*} \frac{1}{\langle q^*, p \rangle} \partial f(x) \subset \partial \psi(x, y)_p,$$

$$(2) \bigcup_{q^* \in K_2^*} \frac{1}{\langle q^*, p \rangle} \partial f(x) = \partial_{\omega} \psi(x, y)_p,$$

其中 $\partial f(x) = \partial f(x, \langle q^*, y \rangle)_1$,

$$K_1^* = \{q^* \in \text{int}K^* \mid \|q^*\|_* = 1\},$$

$$K_2^* = \{q^* \in K^* \setminus \{0\} \mid \|q^*\|_* = 1\}.$$

证明 由定理 6.2.14 和定理 6.2.15 可得证. \square

我们可以减弱定理 6.2.14 和定理 6.2.16 的条件, 分别给出如下结果.

定理 6.2.17 设 $S \subset \mathcal{B}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{B}}$ 是集值映射, $x \in S$, $y \in \psi(x)$, $p \in \text{int}K$, K^* 是 K 的对偶锥. 若 ψ 在 S 上是 K -凸的, 对任意的 $x' \in S$, $\psi(x')$ 是非空紧集, 则 f 在 S 上是

凸的.

(1) 再若存在 $q^* \in \text{int}K^*$ 使得 $f(x) = \langle q^*, y \rangle$, 则

$$\{\langle q^*, p \rangle\}^{-1} \partial f(x, \langle q^*, y \rangle)_1 \subset \partial \psi(x, y)_p.$$

(2) 再若存在 $q^* \in K^* \setminus \{0\}$ 使得 $f(x) = \langle q^*, y \rangle$, 则

$$\{\langle q^*, p \rangle\}^{-1} \partial f(x, \langle q^*, y \rangle)_1 \subset \partial_w \psi(x, y)_p.$$

$$(3) \partial_w(x, y)_p \subset \bigcup_{q^* \in K^* \setminus \{0\}} \frac{1}{\langle q^*, p \rangle} \partial f(x, \langle q^*, y \rangle)_1.$$

证明 类似于定理 6.2.14 的证明可知, f 在 S 上是凸的, 并且 $\partial f(x) = \partial \theta(x, \langle q^*, y \rangle)_1$.

(1) 当 $q^* \in \text{int}K^*$, $x^* \in \{\langle q^*, p \rangle\}^{-1} \partial \theta(x, \langle q^*, y \rangle)_1$, 类似于定理 6.2.13 的(1)的证明, 可以推知 $x^* \in \partial \psi(x, y)_p$.

(2) 与(1)的证明类似可得到.

(3) 由定理 6.2.12 知当 $x^* \in \partial_w \psi(x, y)_p$ 时, 则存在 $q^* \in K^* \setminus \{0\}$ 使得

$$\begin{aligned} \langle q^*, y' - y \rangle &\geq \langle x^*, x' - x \rangle \langle q^*, p \rangle \\ \forall x' \in S, \forall y' \in \psi(x'). \end{aligned} \quad (6.2.21)$$

设集值函数 $\theta(x) = \langle q^*, \psi(x) \rangle$, 在(6.2.21)中令 $x' = x$ 得

$$\langle q^*, y' \rangle \geq \langle q^*, y \rangle, \quad \forall y' \in \psi(x'),$$

故有 $f(x) = \langle q^*, y \rangle$. 再由(6.2.21)得到

$$f(x') - f(x) \geq \langle \langle q^*, p \rangle x^*, x' - x \rangle \quad \forall x' \in S,$$

所以 $\langle q^*, p \rangle x^* \in \partial f(x, \langle q^*, y \rangle)_1$, 这就证明了(3). \square

定理 6.2.18 设 \mathcal{B} 是 Banach 空间, $S \subset \mathcal{B}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{B}}$ 是集值映射, $x \in S, y \in \psi(x), p \in \text{int}K, K^*$ 是 K 的对偶锥. 若 ψ 在 S 上是 K -凸的, 对任意的 $x' \in S, \psi(x')$ 是非空紧集, 则

$$\partial_w \phi(x, y) \subset \bigcup_{q^* \in K_2^*} \frac{1}{\langle q^*, p \rangle} \partial f(x, \langle q^*, y \rangle)_p,$$

其中 $K_2^* = \{q^* \in K^* \setminus \{0\} \mid \|q^*\|_* = 1\}$.

证明 与定理 6.2.17 中(3)的证明类似. \square

定理 6.2.17 提供了我们求锥次梯度和锥弱次梯度的途径.

例 6.2.1 设集合 $S \subset \mathcal{X}$ 非空, 对任意的 $x \in S$, $\phi(x) = T \subset \mathcal{Y}$ 是非空紧集. 考虑集值映射 $\phi: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi(x)$. 这时, 对任何 $q^* \in K^* \setminus \{0\}$, $f(x) = \inf_{y \in T} \langle q^*, y \rangle$ 是一个常数, 所以 $\partial f(x) = \{0\}$. 由定理 6.2.17 和定理 6.2.18, 得 $\partial \phi(x)_p = \{0\} (\forall x \in S)$.

例 6.2.2 设集合 $S \subset \mathcal{X}$ 非空, $\phi = f: S \rightarrow R^m$ 是实向量函数, 取 $K = R_+^m$, 并且 f 是 R_+^m -凸的可微函数. 记 $f(x) = (f_1(x), \dots, f_m(x))^T \in R^m (x \in S)$, 令 $p = (1, \dots, 1)^T \in R^m$, $q^* = (q_1, \dots, q_m)^T \in R_+^m \setminus \{0\}$, $\langle q^*, p \rangle = \sum_{i=1}^m q_i$. 设 $x \in S$, $y = f(x)$, $g(x) = \langle q^*, y \rangle$, 则有

$$\partial g(x, \langle q^*, y \rangle)_1 = \sum_{i=1}^m q_i \nabla f_i(x).$$

由定理 6.2.17 得到

$$\partial f(x, y)_p \supset \bigcup_{q^* \in \text{int} K^*} \frac{1}{\sum_{i=1}^m q_i} \sum_{i=1}^m q_i \nabla f_i(x)$$

和

$$\partial_w f(x, y)_p = \bigcup_{q^* \in K^* \setminus \{0\}} \frac{1}{\sum_{i=1}^m q_i} \sum_{i=1}^m q_i \nabla f_i(x),$$

即

$$\partial f(x)_p \supset \left\{ \sum_{i=1}^m q_i \nabla f_i(x) \mid \sum_{i=1}^m q_i = 1, q_i \in (0, 1) \right\}$$

和

$$\partial_w f(x)_p = \left\{ \sum_{i=1}^m q_i \nabla f_i(x) \mid \sum_{i=1}^m q_i = 1, q_i \in [0, 1] \right\}.$$

由此可知道, f 的分量梯度的凸组合构成了 f 的 K -次微分.

例 6.2.3 考虑实值函数 $g: S \times Y \rightarrow R$, 其中 S 和 Y 分别是 Banach 空间 \mathcal{B} 和 \mathcal{L} 中的非空集合. 作集值函数 $\psi: S \rightarrow 2^R$, $x \mapsto \psi(x) = \{t \mid t = g(x, y), y \in Y\}$, 设 $f(x) = \inf_{y \in Y} g(x, y)$ ($\forall x \in S$), 令 $p = 1$, $q^* = 1$, 且有 $f(x) = \inf \psi(x)$. 如果 g 在 $S \times Y$ 上是下半连续的正常凸函数, 并且当 $x \in S$, $y \in Y$ 时有 $f(x) = g(x, y)$, 则

$$x^* \in \partial f(x) \iff (x^*, 0) \in \partial \psi(x, y)_1.$$

设 $t = g(x, y)$, 由定理 6.2.17 得到

$$\partial \psi(x, t)_1 = \partial_w \psi(x, t)_1 = \partial f(x).$$

例 6.2.4 考虑实值向量函数 $f: S \times Y \rightarrow R^m$, 其中 S 和 Y 分别是 Banach 空间 \mathcal{B} 和 \mathcal{L} 中的非空集合. 设 $p = (1, \dots, 1)^T \in R^m$, $q^* \in R_+^m \setminus \{0\}$, 函数 $g: S \times Y \rightarrow R$, $g(x, y) = \langle q^*, f(x, y) \rangle$ ($\forall (x, y) \in S \times Y$). 作集值函数 $\psi: S \rightarrow 2^{R^m}$, $x \mapsto \psi(x) = \{r \mid r = f(x, y), y \in Y\}$. 令

$$\bar{f}(x) = \inf_{y \in Y} g(x, y) = \inf_{y \in Y} \langle q^*, f(x, y) \rangle,$$

设 f 在 $S \times Y$ 上是下半连续的正常凸函数, 由例 6.2.3 和定理 6.2.17 可得

$$x^* \in \partial \bar{f}(x) \iff (x^*, 0) \in \partial g(x, y) = \sum_{i=1}^m q_i^* \partial f_i(x, y),$$

其中 $q^* = (q_1^*, \dots, q_m^*)^T \in R_+^m \setminus \{0\}$, $f = (f_1, \dots, f_m)^T \in R^m$. 此外, 有

$$\begin{aligned}\partial\phi(x, r)_p &\supset \left\{ x^* \in \mathcal{X}^* \mid (x^*, \theta) \in \sum_{i=1}^m q_i^* \partial f_i(x, y), \right. \\ &\quad \left. \sum_{i=1}^m q_i^* = 1, q_i^* \in (0, 1), i = 1, \dots, m \right\}, \\ \partial_w\phi(x, r)_p &= \left\{ x^* \in \mathcal{X}^* \mid (x^*, \theta) \in \sum_{i=1}^m q_i^* \partial f_i(x, y), \right. \\ &\quad \left. \sum_{i=1}^m q_i^* = 1, q_i^* \in [0, 1], i = 1, \dots, m \right\}.\end{aligned}$$

注意, 上式成立的必要条件是: 存在 q^* 使得 $\bar{f}(x) := \langle q^*, r \rangle = g(x, y)$.

§ 6.3 基本特性

现在讨论锥弱次微分及其相应映射的若干基本特性, 如凸性、闭性、半连续性、连通性和次单调性等.

若无特别说明, 本节均设 $S \subset \mathcal{X}$ 是非空凸集, $\phi: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi(x)$ 是集值映射, 对任意的 $x \in S$ 而言, 集合 $\phi(x) \subset \mathcal{Y}$ 非空, $K \subset \mathcal{Y}$ 是内部非空的尖闭凸锥.

先介绍锥弱次微分的凸性和闭性.

定理 6.3.1 设 $x \in S$, $y \in \phi(x)$, $p \in \text{int}K$. 若 ϕ 在 S 上是 K -凸的, 则 $\partial_w\phi(x, y)_p$ 是凸集.

证明 设 $x_1^*, x_2^* \in \partial_w\phi(x, y)_p$, 对任意的 $\lambda \in (0, 1)$, 记 $x_\lambda^* = \lambda x_1^* + (1 - \lambda)x_2^*$. 为证 $x_\lambda^* \in \partial_w\phi(x, y)_p$, 用反证法, 假设 $x_\lambda^* \notin \partial_w\phi(x, y)_p$, 则按定义 6.1.3 可知, 存在 $\bar{x} \in S$, $\bar{y} \in \phi(\bar{x})$, $(\bar{x}, \bar{y}) \neq (x, y)$, 有

$$y - \langle x_\lambda^*, x \rangle p - (\bar{y} - \langle x_\lambda^*, \bar{x} \rangle p) \in \text{int}K. \quad (6.3.1)$$

由 $x_1^*, x_2^* \in \partial_w\phi(x, y)_p$, 根据定理 6.2.12 知, 存在 $q_1^*, q_2^* \in K^* \setminus \{\theta\}$ 使

$$\langle q_1^*, \bar{y} - y \rangle \geq \langle x_1^*, \bar{x} - x \rangle \langle q_1^*, p \rangle, \quad (6.3.2)$$

$$\langle q_2^*, \bar{y} - y \rangle \geq \langle x_2^*, \bar{x} - x \rangle \langle q_2^*, p \rangle. \quad (6.3.3)$$

从(6.3.1)有

$$\langle q_1^*, y - \bar{y} \rangle + \langle x_1^*, \bar{x} - x \rangle p > 0, \quad (6.3.4)$$

$$\langle q_2^*, y - \bar{y} \rangle + \langle x_2^*, \bar{x} - x \rangle p > 0. \quad (6.3.5)$$

由(6.3.2) + (6.3.4) 和(6.3.3) + (6.3.5), 分别可得

$$\langle x_1^*, \bar{x} - x \rangle > \langle x_1^*, \bar{x} - x \rangle \quad (6.3.6)$$

和

$$\langle x_1^*, \bar{x} - x \rangle > \langle x_2^*, \bar{x} - x \rangle. \quad (6.3.7)$$

作 $\lambda(6.3.6) + (1 - \lambda) \times (6.3.7)$, 则得到

$$\langle x_1^*, \bar{x} - x \rangle > \langle x_1^*, \bar{x} - x \rangle,$$

导致矛盾. \square

定理 6.3.2 设 $x \in S$, $y \in \phi(x)$, $p \in \text{int}K$, ϕ 在 S 上是 K -凸的.

(1) 若 $\partial_w \phi(x, y)_p \neq \emptyset$, 则 $\partial_w \phi(x, y)_p$ 是闭集.

(2) 若 $\phi(x)$ 是紧集, $\partial_w \phi(x)_p \neq \emptyset$, 则 $\partial_w \phi(x)_p$ 是闭集.

证明 (1) 为证 $\partial_w \phi(x, y)_p$ 是闭集, 反之设点列 $\{x_k^*\} \subset \partial_w \phi(x, y)_p$, $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$), 但 $x^* \notin \partial_w \phi(x, y)_p$, 则从后者由定义 6.1.3 可知, 存在 $\bar{x} \in S$ 和 $\bar{y} \in \phi(\bar{x})$ 使得

$$y - \bar{y} + \langle x^*, \bar{x} - x \rangle p \in \text{int}K.$$

因为 $y - \bar{y} + \langle x_k^*, \bar{x} - x \rangle p \rightarrow y - \bar{y} + \langle x^*, \bar{x} - x \rangle p$, 所以存在充分大的 k' 使得

$$y - \bar{y} + \langle x_{k'}^*, \bar{x} - x \rangle p \in \text{int}K.$$

由此, 根据定义 6.1.3 知 $x_{k'}^* \in \partial_w \phi(x, y)_p$, 它与已设矛盾.

(2) 设点列 $\{x_k^*\} \subset \partial_w \phi(x)_p$, $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$), 则存在 $\{y^k\}$

$\subset \mathcal{Y}$, $y^k \in \psi(\bar{x})$, 使 $x_k^* \in \partial_w \psi(x, y)_p$. 由于 $\psi(x)$ 是紧集, 因而 $\{y^k\}$ 存在收敛的子列, 不妨即设 $y^k \rightarrow y$ ($k \rightarrow \infty$), $y \in \psi(x)$. 现假设 $x^* \notin \partial_w \psi(x, y)_p$, 则由定义 6.1.3 的(2)可知, 存在 $\bar{x} \in S$ 和 $\bar{y} \in \psi(\bar{x})$ 使得

$$y - \bar{y} + \langle x^*, \bar{x} - x \rangle p \in \text{int}K. \quad (6.3.8)$$

从 $x_k^* \rightarrow x^*$ 和 $y^k \rightarrow y$ 有

$$y^k - \bar{y} + \langle x_k^*, \bar{x} - x \rangle p \rightarrow y - \bar{y} + \langle x^*, \bar{x} - x \rangle p,$$

于是由(6.3.8)知, 存在充分大的 k' 使

$$y^{k'} - \bar{y} + \langle x_{k'}^*, \bar{x} - x \rangle p \in \text{int}K.$$

据此, 得到 $x_{k'}^* \notin \partial_w \psi(x, y_{k'})_p$, 它与已设矛盾, 从而有 $x^* \in \partial_w \psi(x, y)_p \subset \partial_w \psi(x)_p$. \square

以下考虑锥弱次微分映射 $\partial_w \psi_p: S \rightarrow 2^{\mathcal{W}^*}$, $x \mapsto \partial_w \psi(x)_p$ 的上半连续性.

定理 6.3.3 设 $x \in S$, $y \in \psi(x)$, $p \in \text{int}K$, 则集值映射 $\psi: S \rightarrow 2^{\mathcal{Y}}$ 在 (x, y) 处的 K -弱次微分映射 $\partial_w \psi_p$ 在点 (x, y) 处是上半连续的.

证明 设点列 $\{x^k\} \subset S$, $x^k \rightarrow x$ ($k \rightarrow \infty$), $\{y^k\} \subset \mathcal{Y}$, $y^k \in \psi(x^k)$, $y^k \rightarrow y$ ($k \rightarrow \infty$), 有 $x_k^* \in \partial_w \psi(x^k, y^k)$ 和 $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$). 按定义 3.3.1 的(1), 反之假设 $x^* \notin \partial_w \psi(x, y)_p$, 则由定义 6.1.3 的(2)可知存在 $\bar{x} \in S$, $\bar{y} \in \psi(\bar{x})$ 使得

$$y - \bar{y} + \langle x^*, \bar{x} - x \rangle p \in \text{int}K. \quad (6.3.9)$$

从 $x^k \rightarrow x$, $y^k \rightarrow y$ 和 $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$), 有

$$y^k - \bar{y} + \langle x_k^*, \bar{x} - x^k \rangle p \rightarrow y - \bar{y} + \langle x^*, \bar{x} - x \rangle p.$$

于是, 由(6.3.9)得知, 存在充分大的 k' 使

$$y^{k'} - \bar{y} + \langle x_{k'}^*, \bar{x} - x^{k'} \rangle p \in \text{int}K.$$

据此, 得到 $x_k^* \notin \partial_w \phi(x^k, y^k)_p$, 导致与已设矛盾. \square

定义 6.3.1 设点 $x^0 \in S$, 集合 $\phi(x) \subset \mathscr{Y}$ ($x \in S$). 若存在点 x^0 的邻域 $U(x^0)$ 使 $\text{cl} \bigcap_{x \in U(x^0)} \phi(x) \subset \mathscr{Y}$ 是紧集, 则称 ϕ 在点 x^0 附近是一致紧的.

定理 6.3.4 设 $x \in S$, $p \in \text{int}K$. 若集值映射 $\phi: S \rightarrow 2^*$ 在 x 处是上半连续的, 并且在 x 附近是一致紧的, 则 K -弱次微分映射 $\partial_w \phi_p: S \rightarrow 2^{**}$, $x \mapsto \partial_w \phi(x)_p$ 在 x 处是上半连续的.

证明 设点列 $\{x^k\} \subset S$, $x^k \rightarrow x$ ($k \rightarrow \infty$), $x_k^* \in \partial_w \phi(x^k)_p$ 和 $x_k^* \rightarrow x^*$ ($k \rightarrow \infty$). 由定义 3.3.1 的(1), 只要证明 $x^* \in \partial_w \phi(x)_p$. 由于 $x_k^* \in \partial_w \phi(x^k)_p$, 存在 $y^k \in \phi(x^k)$ 使得

$$x_k^* \in \partial_w \phi(x^k, y^k)_p.$$

因为 $x^k \rightarrow x$ ($k \rightarrow \infty$) 和 ϕ 在 x 附近是一致紧的, 故知 $\{y^k\}$ 存在收敛子列. 不妨设 $y^k \rightarrow y$, 由 ϕ 在 x 处是上半连续的, 按定义 3.3.1 的(1)可知 $y \in \phi(x)$. 于是, 根据定理 6.3.4 得到 $x^* \in \partial_w \phi(x, y)_p$, 从而 $x^* \in \partial_w \phi(x)_p$. \square

下面讨论锥弱次微分的连通性.

定义 6.3.2 设 S 是任意非空集合. 若不存在两个集合 $S_1 \neq \emptyset$ 和 $S_2 \neq \emptyset$, 使得

$$S = S_1 \cup S_2,$$

其中 $\text{cl}S_1 \cap S_2 = S_1 \cap \text{cl}S_2 = \emptyset$, 则称 S 是连通集, 或 S 是连通的. 不是连通的集合称为是可分离的.

引理 6.3.5 设 \mathscr{V} 是线性空间. 若 $S \subset \mathscr{V}$ 是非空凸集, 则 S 是连通集.

证明 反之, 假设 S 是可分离的, 则由定义 6.3.2 可知, 存在集合 $S_1 \neq \emptyset$ 和 $S_2 \neq \emptyset$, 并且

$$\text{cl}S_1 \cap S_2 = S_1 \cap \text{cl}S_2 = \emptyset, \quad (6.3.10)$$

使得

$$S = S_1 \cup S_2. \quad (6.3.11)$$

设 $v^1 \in S_1$, $v^2 \in S_2$, 由 (6.3.11) 注意到 S 是凸集, 故对任意的 $\lambda \in (0, 1)$ 有 $\lambda v^1 + (1 - \lambda)v^2 \in S$, 从而对于线段 $[v^1, v^2]$ 有

$$[v^1, v^2] \subset S. \quad (6.3.12)$$

记 $S_3 = [v^1, v^2] \cap S_1$ 和 $S_4 = [v^1, v^2] \cap S_2$, 因为 $v^1 \in S_1$ 和 $v^2 \in S_2$, 故有 $S_3 \neq \emptyset$ 和 $S_4 \neq \emptyset$. 由 (6.3.11) 和 (6.3.12), 我们有

$$\begin{aligned} S_3 \cup S_4 &= ([v^1, v^2] \cap S_1) \cup ([v^1, v^2] \cap S_2) \\ &= [v^1, v^2] \cap (S_1 \cup S_2) = [v^1, v^2] \cap S \\ &= [v^1, v^2]. \end{aligned}$$

又由 (6.3.10) 可知, 有

$$\begin{aligned} \text{cl}S_3 \cap S_4 &= \text{cl}([v^1, v^2] \cap S_1) \cap ([v^1, v^2] \cap S_2) \\ &= [v^1, v^2] \cap \text{cl}S_1 \cap S_2 = \emptyset. \end{aligned}$$

同理, 有 $S_3 \cap \text{cl}S_4 = \emptyset$. 由此, 从以上所得根据定义 6.3.2 导致线段 $[v^1, v^2]$ 是可分离的, 但这是不可能的, 从而定理得证. \square

引理 6.3.6 设 $S \subset \mathcal{X}$ 是连通集, 集值映射 $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是上半连续的. 若对任意的 $x \in S$, $\phi(x) \neq \emptyset$ 是连通集, 则 $\bigcup_{x \in S} \phi(x)$ 是连通集.

证明 见文献[27]的命题 2.5.6. \square

引理 6.3.7 设 $Y \subset \mathcal{Y}$ 是非空紧凸集, 则 $\mathcal{C}_w(Y, K)$ 是连通集.

证明 见文献[27]的定理 3.5.12. \square

定理 6.3.8 设 $x \in S$, $y \in \phi(x)$, $\phi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $p \in \text{int}K$.

(1) 若 $\partial_w \phi(x, y)_p \neq \emptyset$, 则 $\partial_w \phi(x, y)_p$ 是连通集.

(2) 若 $\phi(x)$ 是紧凸集, ϕ 在 S 上是上半连续的, 则 $\partial_w \phi(x)_p$ 是连通集.

证明 (1) 由定理 6.3.2 知 $\partial_w \phi(x, y)_p$ 是凸集, 再由引理

6.3.6 即知 $\partial_w \phi(x, y)_p$ 是连通集.

(2) 因 $\phi(x)$ 是紧集, 故也是 K -半紧的, 由定理 6.2.10 可知, 对任意的 $y \in \mathcal{E}_w(\phi(x), K)$ 有 $\partial_w \phi(x, y)_p \neq \emptyset$. 于是, 据定理 6.1.1 的(3)和定义 6.1.3 的(2)得知

$$\begin{aligned} \partial_w \phi(x)_p &= \{x^* \in \partial_w \phi(x, y)_p \mid y \in \mathcal{E}_w(\phi(x), K)\} \\ &= \bigcup_{y \in \mathcal{E}_w(\phi(x), K)} \partial_w \phi(x, y)_p. \end{aligned} \quad (6.3.13)$$

又因为 $\phi(x)$ 是紧凸集, 故由定理 6.3.8 知 $\mathcal{E}_w(\phi(x), K)$ 是连通集. 往下证明 K -弱次微分映射 $\partial_w \phi_p$ 关于 y 在 y 处是上半连续的. 事实上, 设对任意的 $y \in \mathcal{E}_w(\phi(x), K)$, 考虑点列

$$\{y^k\} \subset \mathcal{E}_w, \quad y^k \rightarrow y \quad (k \rightarrow \infty), \quad x_i^* \in \partial_w \phi(x, y^k)_p, \quad x_i^* \rightarrow x^*.$$

根据定理 6.3.4 即知 $\partial_w \phi(x, z)_p$ 在 y 处是上半连续的. 最后, 利用引理 6.3.7 和(6.3.13), 使得 $\partial_w \phi(x)_p$ 是连通集. \square

最后, 讨论锥弱次微分的次单调性.

定理 6.3.9 设 $x^1, x^2 \in S$, $y^1 \in \phi(x^1)$, $y^2 \in \phi(x^2)$. 若 $x_1^* \in \partial_w \phi(x^1, y^1)_p$ 和 $x_2^* \in \partial_w \phi(x^2, y^2)_p$, 则 $\langle x_1^* - x_2^*, x^1 - x^2 \rangle \geq 0$.

证明 若 $x_1^* = x_2^*$ 或者 $x^1 = x^2$, 则定理结论显然成立. 以下设 $x_1^* \neq x_2^*$, $x^1 \neq x^2$. 用反证法, 假设 $\delta = \langle x_1^* - x_2^*, x^1 - x^2 \rangle < 0$. 记

$$a^1 = y^1 - y^2 - \langle x_1^*, x^1 - x^2 \rangle p,$$

$$a^2 = y^2 - y^1 - \langle x_2^*, x^2 - x^1 \rangle p.$$

对任意的 $\lambda \in [0, 1]$, 令 $F(\lambda) = \lambda a^1 + (1 - \lambda) a^2$, 则有

$$F(\lambda) = (2\lambda - 1) a^1 - (1 - \lambda) \delta p, \quad (6.3.14)$$

或即

$$F(\lambda) = (1 - 2\lambda) a^2 - \lambda \delta p. \quad (6.3.15)$$

作集合

$$E = \{y \in \mathcal{E}_w \mid y = F(\lambda) + (1/2) \delta p, \lambda \in [0, 1]\}.$$

显然 E 是凸集, 往下证明

$$E \cap \text{int}K = \emptyset. \quad (6.3.16)$$

事实上, 反之, 若存在 $q \in E \cap \text{int}K$, 由 E 的定义有

$$F(\lambda) + (1/2)\delta p = q, \text{ 对某个 } \lambda \in [0, 1]. \quad (6.3.17)$$

当 $\lambda = 1/2$ 时, 得 $q = \theta \in \text{int}K$, 导致矛盾. 当 $0 \leq \lambda < 1/2$ 时, 由

$$(6.3.15) \text{ 和 } (6.3.17) \text{ 得 } a^2 = \frac{2q - (2\lambda - 1)\delta p}{2(1 - 2\lambda)} \in \text{int}K, \text{ 即}$$

$$y^2 - y^1 - \langle x_2^*, x^2 - x^1 \rangle p \in \text{int}K.$$

由此, 根据定义 6.1.3 的(2)知 $x_2^* \notin \partial_{\omega}\psi(x^2, y^2)_p$, 这与已设矛盾. 当 $1/2 < \lambda \leq 1$ 时, 由(6.3.14)和(6.3.17)得 $a^1 = \frac{2q - (1 - 2\lambda)\delta p}{2(2\lambda - 1)} \in \text{int}K$, 即

$$y^1 - y^2 - \langle x_1^*, x^1 - x^2 \rangle p \in \text{int}K.$$

于是, 由定义 6.1.3 的(2)知 $x_1^* \notin \partial_{\omega}\psi(x^1, y^1)_p$, 也与已设矛盾.

现在, 从(6.3.16)由定理 1.3.7 的(1)可知, 存在 $y^* \in \mathcal{Y}^* \setminus \{\theta\}$ 使得

$$\langle y^*, q \rangle \geq \langle y^*, y \rangle \quad \forall q \in K, \forall y \in E. \quad (6.3.18)$$

因为 $\theta \in E$, 所以从(6.3.18)得 $\langle y^*, q \rangle \geq 0$ ($\forall q \in K$), 从而 $y^* \in K^* \setminus \{\theta\}$. 由 $p \in \text{int}K$ 和定理 1.4.11 的(1)得到 $\langle y^*, p \rangle > 0$. 将 $a^1 + (1/2)\delta p, a^2 + (1/2)\delta p \in E$ 代入(6.3.18), 我们有

$$\langle y^*, a^1 \rangle + (1/2)\delta p \leq 0, \quad (6.3.19)$$

$$\langle y^*, a^2 \rangle + (1/2)\delta p \leq 0. \quad (6.3.20)$$

由(6.3.18)+(6.3.19)得 $\langle y^*, -\delta p \rangle + (1/2)\delta p \leq 0$, 即 $\delta \langle y^*, p \rangle \geq 0$, 因此推知 $\delta \geq 0$, 导致与假设矛盾. \square

定理 6.3.10 设 $x \in S, y \in \psi(x), p^1, p^2 \in \text{int}K$. 若 $x^* \in \partial_{\omega}\psi(x, y)_{p^1}$, 则存在 $q^* \in K^* \setminus \{\theta\}$ 使得 $\alpha x^* \in \partial_{\omega}\psi(x, y)_{p^2}$, 其中

$$\alpha = \frac{\langle q^*, p^1 \rangle}{\langle q^*, p^2 \rangle}.$$

证明 由定理 6.2.12 知, 存在 $q^* \in K^* \setminus \{0\}$ 使得

$$\langle q^*, y' - y \rangle \geq \langle x^*, x' - x \rangle \langle q^*, p^1 \rangle$$

$$\forall x' \in S, \forall y' \in \phi(x').$$

令 $\alpha = \frac{\langle q^*, p^1 \rangle}{\langle q^*, p^2 \rangle}$, 注意 $\langle q^*, p^1 \rangle > 0$ 和 $\langle q^*, p^2 \rangle > 0$ 有 $\alpha > 0$, 故得

$$\langle q^*, y' - y \rangle \geq \langle \alpha x^*, x' - x \rangle \langle q^*, p^2 \rangle$$

$$\forall x' \in S, \forall y' \in \phi(x').$$

由定理 6.2.11 的(2), 即得 $\alpha x^* \in \partial_w \phi(x, y)_{p^2}$. \square

注 6.3.1 定理 6.3.11 表明了 $\partial_w \phi(x, y)_{p^1}$ 与 $\partial_w \phi(x, y)_{p^2}$ 中的元素仅相差一个倍数. 因此, 只要求得关于某个 p 的锥弱次微分就可得到所有的结果. 从这一意义上, 我们说 $\partial_w \phi(x, y)_p$ 与 p 是无关的.

定理 6.3.11 设实值向量函数 $f: S \rightarrow R^m$ 是 R_+^m -凸的, $f(x) = [f_1(x), \dots, f_m(x)]^T$ 在 S 上每一点处有定义, 令 $p = (1, \dots, 1)^T \in R^m$. 若 f 在 S 上是连续的, $x \in S$, 则

$$\partial_w f(x)_p = \left\{ \sum_{i=1}^m \mu_i \partial f_i(x) \mid \sum_{i=1}^m \mu_i = 1, \mu_i \geq 0, 1, \dots, m \right\},$$

其中 $\partial f_i(x)$ 是凸函数的次微分.

证明 由定理 6.2.13 的(2)和凸函数次微分的可加性, 可得

$$\begin{aligned} \partial_w f(x)_p &= \bigcup_{q^* \in K^* \setminus \{0\}} \frac{1}{\langle q^*, p \rangle} \partial \theta(x, \langle q^*, y \rangle)_1 \\ &= \bigcup_{q^* \in K^* \setminus \{0\}} \left(\sum_{i=1}^m q_i^* \right)^{-1} \sum_{i=1}^m q_i^* \partial f_i(x) \end{aligned}$$

$$= \left\{ \sum_{i=1}^m \mu_i \partial f_i(x) \mid \sum_{i=1}^m \mu_i = 1, \mu_i \geq 0, i = 1, \dots, m \right\}.$$

于是, 定理得证. \square

显然, 例 6.2.2 可看成是定理 6.3.12 的推论.

例 6.3.1 考虑实值向量函数 $f: R^2 \rightarrow R^2$, $f(x, y) = (|x|, |y|)^T$, $(x, y)^T \in R^2$. 设 $f_i(x) = |x|$, $x \in R$ ($i = 1, 2$), 则有 $\partial f_i(0) = [-1, 1]$, 根据定理 6.3.12 得

$$\begin{aligned} \partial_w f(0, 0)_p &= \{(\mu_1 a, \mu_2 b)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \\ &\quad \mu_1, \mu_2 \geq 0, a, b \in [-1, 1]\}. \end{aligned}$$

此外, 当 $x > 0, y = 0$ 时, 有

$$\begin{aligned} \partial_w f(x, 0)_p &= \{(\mu_1, \mu_2 b)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \\ &\quad \mu_1, \mu_2 \geq 0, b \in [-1, 1]\}; \end{aligned}$$

当 $x = 0, y > 0$ 时, 有

$$\begin{aligned} \partial_w f(0, y)_p &= \{(\mu_1 a, \mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \\ &\quad \mu_1, \mu_2 \geq 0, a \in [-1, 1]\}; \end{aligned}$$

当 $x > 0, y > 0$ 时, 有

$$\partial_w f(x, y)_p = \{(\mu_1, \mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\};$$

当 $x > 0, y < 0$ 时, 有

$$\partial_w f(x, y)_p = \{(\mu_1, -\mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\};$$

当 $x < 0, y > 0$ 时, 有

$$\partial_w f(x, y)_p = \{(-\mu_1, \mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\};$$

当 $x < 0, y < 0$ 时, 有

$$\partial_w f(x, y)_p = \{(-\mu_1, -\mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\}.$$

$\partial_w f(x, y)_p$ 的 9 种情况如表 6.3.1 所示.

表 6.3.1

x	$x = 0$			$x > 0$			$x < 0$		
y	$y = 0$	$y > 0$	$y < 0$	$y = 0$	$y > 0$	$y < 0$	$y = 0$	$y > 0$	$y < 0$
$\partial_w f(x, y)_p$	$ABCD$	ABC	ACD	BCD	BC	CD	ABD	AB	AD

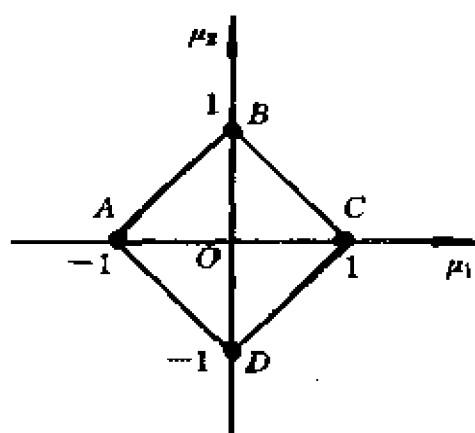


图 6.3.1

图 6.3.1 中的 $ABCD$ 表示由直线 AB 、 BC 、 CD 和 AD 围成的方形区域; ABC 表示由直线 AB 、 BC 和 CA 围成的三角形区域; AB 表示直线. $\partial_w f(x, y)_p$ 与相应的区域对应.

§ 6.4 运算性质

本节给出锥弱次微分的基本运算性质,如次可加性和复合映射的链式法则等.

设集合 $S \subset \mathcal{X}$, $\phi: \mathcal{X} \rightarrow 2^Y$, $x \mapsto \phi(x)$ 是集值映射, $K \subset Y$ 是内部非空的尖闭凸锥.

定理 6.4.1 设 $S \subset \mathcal{X}$ 是非空凸集, 集值映射 $\phi: S \rightarrow 2^Y$ 是 K -凸的, $x \in S$, $y \in \phi(x)$, $p \in \text{int}K$. 若 $\alpha > 0$, 则

$$\partial_w(\alpha\phi)(x, y)_p = \alpha\partial_w\phi(x, y)_p. \quad (6.4.1)$$

证明 设 $\alpha = 0$, 则 (6.4.1) 显然成立. 设 $\alpha > 0$, 任取 $x^* \in \partial_w(\alpha\psi)(x, y)_p$, 由定理 6.2.12 知存在 $q^* \in K^* \setminus \{0\}$ 使得

$$\langle q^*, \alpha y' - \alpha y \rangle \geq \langle x^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall \alpha y' \in \alpha\psi(x'),$$

即

$$\langle q^*, y' - y \rangle \geq \langle x^*/\alpha, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

由此, 根据定理 6.2.11 的(2)得知 $x^* \in \alpha \partial_w \psi(x, y)_p$. 推导的逆过程也是成立的. 定理证毕. \square

定理 6.4.2 设 $S \subset \mathcal{X}$ 是内部非空的凸集, 集合 $\psi_1(x)$, $\psi_2(x) \subset \mathcal{Y}$ ($x \in S$) 非空,

$$(\psi_1 + \psi_2)(x) = \{y^1 + y^2 \mid y^1 \in \psi_1(x), y^2 \in \psi_2(x), x \in S\},$$

$$x \in \text{int}S, y^1 \in \psi_1(x), y^2 \in \psi_2(x).$$

对于集值映射 $\psi_1: S \rightarrow 2^{\mathcal{Y}}, x \mapsto \psi_1(x)$; $\psi_2: S \rightarrow 2^{\mathcal{Y}}, x \mapsto \psi_2(x)$; 以及 $\psi_1 + \psi_2: S \rightarrow 2^{\mathcal{Y}}, x \mapsto (\psi_1 + \psi_2)(x)$, 若 ψ_1 和 ψ_2 在 S 上是 K -凸的, 并且 ψ_1 或 ψ_2 在点 x 处是 K -有界逆开的(或连通的), 则

$$\partial_w(\psi_1 + \psi_2)(x, y^1 + y^2)_p \subset \partial_w \psi_1(x, y^1)_p + \partial_w \psi_2(x, y^2)_p, \quad (6.4.2)$$

$$\partial_w(\psi_1 + \psi_2)(x)_p \subset \partial_w \psi_1(x)_p + \partial_w \psi_2(x)_p. \quad (6.4.3)$$

证明 设 $x^* \in \partial_w(\psi_1 + \psi_2)(x, y)_p$, 不妨设 ψ_2 在 x 处满足 K -有界逆开的. 作集合

$$E = \{(x', y') \in \mathcal{X} \times \mathcal{Y} \mid x' \in S, y' = y^1 - y_1^1 + \langle x^*, x' - x \rangle p - q^1, y_1^1 \in \psi_1(x'), q^1 \in K\},$$

$$F = \{(x', y') \in \mathcal{X} \times \mathcal{Y} \mid x' \in S, y' = y_2^2 - y^2 + q^2,$$

$$y'_2 \in \psi_2(x'), q^2 \in K\}.$$

显然 $E \neq \emptyset$, 根据引理 6.2.1 知 $\text{int}(K\text{-epi}\psi_2) \neq \emptyset$, 从而 $\text{int}F \neq \emptyset$. 由已知条件不难推得 E 和 F 是凸集, 下面证明

$$E \cap \text{int}F = \emptyset. \quad (6.4.4)$$

事实上, 否则, 假若 $E \cap \text{int}F \neq \emptyset$, 则存在 $\bar{x} \in S$, $\bar{y}_1 \in \psi_1(\bar{x})$, $\bar{y}_2 \in \psi_2(\bar{x})$, 以及 $\bar{q}_1, \bar{q}_2 \in K$, 使得

$$\begin{aligned} \bar{y} &= \bar{y}_2 - y^2 + \bar{q}_2 \\ &= y^1 - \bar{y}_1 + \langle x^*, \bar{x} - x \rangle p - \bar{q}_1. \end{aligned} \quad (6.4.5)$$

由于 $(\bar{x}, \bar{y}) \in \text{int}F$, 因而存在充分小的 $\varepsilon > 0$ 有 $(\bar{x}, \bar{y} - \varepsilon p) \in F$. 据此, 有 $y'_2 \in \psi_2(\bar{x})$, $q'_2 \in K$, 使得

$$y'_2 - y^2 + q'_2 = \bar{y}_2 - y^2 + \bar{q}_2 - \varepsilon p.$$

将上式代入(6.4.5), 有

$$y'_2 - y^2 + q'_2 + \varepsilon p = y^1 - \bar{y}_1 + \langle x^*, \bar{x} - x \rangle p - \bar{q}_1,$$

因为 K 是凸锥, 故

$$\begin{aligned} y^1 + y^2 - \langle x^*, x \rangle p - (\bar{y}_1 + y'_2 - \langle x^*, \bar{x} \rangle p) \\ = \bar{q}_1 + \varepsilon p \in \text{int}K. \end{aligned}$$

于是, 由定义 6.1.3 的(2)知 $x^* \notin \partial_w(\psi_1 + \psi_2)(x, y)_p$, 这与已设矛盾.

由(6.4.4)根据定理 1.3.7 的(1), 得知存在 $(x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^*$ 和 $(x^*, y^*) \neq \theta$, 使得

$$\begin{aligned} \langle x^*, x_e \rangle + \langle y^*, y_e \rangle &\geq \beta \geq \langle x^*, x_f \rangle + \langle y^*, y_f \rangle \\ \forall (x_e, y_e) \in E \quad \forall (x_f, y_f) \in F. \end{aligned} \quad (6.4.6)$$

在上式中 $y^* \neq \theta$, 因为否则, 令 $x_e = x'$, $x_f = x$ 代入(6.4.6)得 $\langle x^*, x' - x \rangle \geq 0$ ($\forall x' \in S$). 类似于定理 6.2.4 的证明, 导致与

$x \in \text{int} S$ 矛盾. 由于对任意的 $q \in K$ 有 $(x, -q) \in E$ 和 $(x, q) \in F$, 代入 (6.4.6) 得

$$\langle -y^*, q \rangle > 0 \quad \forall q \in K,$$

因此 $-y^* \in K^* \setminus \{0\}$. 令 $q^* = -y^*$, 由 $p \in \text{int} K$, 根据定理 1.4.11 的 (1) 可知 $\langle q^*, p \rangle > 0$. 又由 $(x, 0) \in E \cap F$, 代入 (6.4.6) 得 $\beta = \langle x^*, x \rangle$. 设

$$x_e = x_f = x', y_e = \bar{y}_1 - y'_1 + \langle x^*, x' - x \rangle p,$$

$$y_f = y'_2 - y^2,$$

将以上式子一同代入 (6.4.6), 得到

$$\begin{aligned} & \langle x^*, x' \rangle + \langle -q^*, y^1 - y'_1 \rangle + \langle x^*, x' - x \rangle p \\ & \geq \langle x^*, x \rangle \geq \langle x^*, x' \rangle + \langle -q^*, y'_2 - y^2 \rangle. \end{aligned}$$

令 $x_2^* = x^* / \langle q^*, p \rangle$, $x_1^* = x^* - x_2^*$, 则有

$$\langle q^*, y'_1 - y^1 \rangle \geq \langle x_1^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y'_1 \in \phi_1(x').$$

$$\langle q^*, y'_2 - y^2 \rangle \geq \langle x_2^*, x' - x \rangle \langle q^*, p \rangle$$

$$\forall x' \in S, \forall y'_2 \in \phi_2(x').$$

据此, 由定理 6.2.11 的 (2) 知 $x_1^* \in \partial \phi_1(x, y^1)_p$, $x_2^* \in \partial \phi_2(x, y^2)_p$, 因而 (6.4.2) 成立.

由 (6.4.2) 和定义 6.1.3 的 (2), 容易推出 (6.4.3) 成立. \square

注 6.4.1 仅当 ϕ_1 和 ϕ_2 是单值凸函数时, (6.4.2) 才成为等式.

注 6.4.2 (6.4.2) 的反包含关系一般不成立, 即

$$\partial_w \phi_1(x, y^1)_p + \partial_w \phi_2(x, y^2)_p \subset \partial_w (\phi_1 + \phi_2)(x, y^1 + y^2)_p$$

不一定成立. 例如, 设 $\mathcal{X} = R$, $\mathcal{Y} = R^2$, $K = R_+^2$, $p = (1, 1)^T$,

$\psi_1(x) = (-x, x)^T$, $\psi_2(x) = (x, -x)^T$, 则 $\psi_1, \psi_2: R \rightarrow R^2$ 是连续的和 R_+^2 -凸的, 并且 $\psi_1(0) = \theta$, $\psi_2(0) = \theta$. 由定理 6.3.12 得

$$\partial_w \psi_1(0)_p = \{-\mu_1 + \mu_2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\},$$

$$\partial_w \psi_2(0)_p = \{\mu_1 - \mu_2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\},$$

故有

$$\partial_w \psi_1(0)_p + \partial_w \psi_2(0)_p = \{2a - 2b \mid 0 \leq a \leq 1, 0 \leq b \leq 1\}.$$

另外, 由于 $(\psi_1 + \psi_2)(x) = (0, 0)^T$ ($\forall x \in R$), 则有

$$\partial_w(\psi_1 + \psi_2)(0)_p = \{\theta\}.$$

推论 6.4.3 设 $S \subset \mathcal{X}$ 是内部非空的开凸集, 集合 $\psi(x) \subset R^m$ ($x \in S, m \geq 1$) 是非空的, $p = (1, \dots, 1)^T \in R^m$. 记

$$\psi(x) = (\psi_1(x), \dots, \psi_m(x))^T \subset R^m,$$

若集值映射 $\psi: S \rightarrow 2^{R^m}$ 是 R_+^m -凸的, $x \in \text{int}S$, $\psi(x)$ 在 x 处是连通的, 则

$$\begin{aligned} \partial_w \psi(x)_p \subset \left\{ \sum_{i=1}^m \mu_i \partial_w \psi_i(x) \mid \sum_{i=1}^m \mu_i = 1, \right. \\ \left. \mu_i \in [0, 1], i = 1, \dots, m \right\}. \end{aligned} \quad (6.4.7)$$

当 ψ 是单值函数时, 上式成为等式.

证明 若 $\partial_w \psi(x)_p = \emptyset$, 则 (6.4.7) 显然成立. 设 $\partial_w \psi(x)_p \neq \emptyset$, 则存在 $y \in \psi(x)$, 使得 $x^* \in \partial_w \psi(x, y)_p$. 由此, 根据定理 6.2.12, 可知存在 $q^* \in R_+^m \setminus \{\theta\}$ 使得

$$\begin{aligned} \langle q^*, y' - y \rangle &\geq \langle x^*, x' - x \rangle \langle q^*, p \rangle \\ \forall x' \in S, \forall y' \in \psi(x'). \end{aligned} \quad (6.4.8)$$

设 $q^* = (q_1^*, \dots, q_m^*)^T$, 记 $\alpha = \langle q^*, p \rangle = \sum_{i=1}^m q_i^*$, 作集值函数

$$F: S \rightarrow 2^R, F(x) = \{f | f = \langle q^*, y \rangle, y \in \psi(x)\},$$

由(6.4.8)得

$$f - f' \geq \langle \alpha x^*, x' - x \rangle, \quad x' \in S, f \in F(x'),$$

其中 $f' = \langle q^*, y \rangle$. 因为 ψ 是 R_+^m -凸的, 则 F 是 R_+^m -凸的, 由定理 6.2.11 得 $\alpha x^* \in \partial_w F(x, f)_1$. 再由 ψ 的分量 $\psi_i (i = 1, \dots, m)$ 在 S 上是 R_+ -凸的, 并且在 x 处是连通的, 根据定理 6.4.1 和定理 6.4.2, 有

$$\partial_w F(x)_1 \subset \sum_{i=1}^m q_i^* \partial_w \psi_i(x)_1,$$

因此

$$x^* \in \sum_{i=1}^m \frac{q_i^*}{\alpha} \partial_w \psi_i(x)_1.$$

令 $\mu_i = q_i^* / \alpha (i = 1, \dots, m)$, 即得(6.4.7).

当 ψ 是单值函数时, 由定理 6.3.12 即知结论成立. \square

最后, 讨论复合映射的求锥弱次微分的问题.

定义 6.4.1 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射. 若对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} & \lambda \psi(x^1) + (1 - \lambda) \psi(x^2) \\ & \subset \psi(\lambda x^1 + (1 - \lambda)x^2) \quad \forall x^1, x^2 \in S, \end{aligned}$$

则称 ψ 是 S 上的凸集值映射, 或 ψ 在 S 上是凸的.

显然, 集值映射 ψ 在 S 上是凸的, 则对任何尖闭凸锥 $K \subset \mathcal{Y}$, 它在 S 上也是 K -凸的.

设 \mathcal{X} , \mathcal{Y} 和 \mathcal{Z} 是局部凸线性拓扑空间, $K_1 \subset \mathcal{Z}$ 是内部非空的尖闭凸锥.

定理 6.4.4 设 $S \subset \mathcal{X}$ 是非空凸集, 集合 $F(x) \subset \mathcal{Y} (\forall x \in S)$ 是非空的, $Y = F(S) = \bigcup_{x \in S} F(x)$, 集合 $G(y) \subset \mathcal{Z} (y \in Y)$ 是非空的. 又设集值映射 $F: S \rightarrow 2^{\mathcal{Y}}$ 在 S 上是凸的, 集值映射 $G: Y \rightarrow$

2^Z 在 Y 上是 K_1 -凸的, $p_1 \in \text{int}K_1$. 考虑集值映射

$$G \circ F: S \rightarrow 2^Z, x \mapsto G \circ F(x) = \bigcup_{y \in F(x)} G(y),$$

设 $x \in S, y \in F(x), z \in G(y), y \in \text{int}Y$. 若 G 在 y 处是 K_1 -有界逆开的或连通的, 则

$$\partial_w G \circ F(x, z)_{p_1} = \{\partial\theta(x, \langle y^*, y \rangle)_1 \mid y^* \in \partial_w G(y, z)_{p_1}\}, \quad (6.4.9)$$

其中集值函数

$$\theta: S \rightarrow 2^R, x \mapsto \theta(x) = \{\langle y^*, y \rangle \mid y \in F(x)\},$$

$\partial\theta(x, \cdot)_1$ 是 θ 在 x 处关于 $p = 1$ 的 R_+ 一次微分.

证明 设 $x^* \in \{\partial\theta(x, \langle y^*, y \rangle)_1 \mid y^* \in \partial_w G(y, z)_{p_1}\}$, 先证明 θ 在 S 上是 R_+ -凸的. 为此对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 任取 $y^1 \in F(x^1), y^2 \in F(x^2)$, 由 F 在 S 上是凸的, 则

$$\lambda y^1 + (1 - \lambda)y^2 \in F(\lambda x^1 + (1 - \lambda)x^2).$$

因为 $\lambda \langle y^*, y^1 \rangle + (1 - \lambda)\langle y^*, y^2 \rangle = \langle y^*, \lambda y^1 + (1 - \lambda)y^2 \rangle$, 所以有

$$\langle y^*, \lambda y^1 + (1 - \lambda)y^2 \rangle \in \langle y^*, F(\lambda x^1 + (1 - \lambda)x^2) \rangle.$$

由此有 $\lambda\theta(x^1) + (1 - \lambda)\theta(x^2) \subset \theta(\lambda x^1 + (1 - \lambda)x^2)$. 再由 F 在 S 上是凸的, 得知 Y 是凸集. 下面证明 $G \circ F$ 在 S 上是 K_1 -凸的. 设对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$, 任取 $z^1 \in G \circ F(x^1), z^2 \in G \circ F(x^2)$, 则存在 $y^1 \in F(x^1)$ 和 $y^2 \in F(x^2)$, 使得 $z^1 \in G(y^1)$ 和 $z^2 \in G(y^2)$. 由于 G 在 Y 上是 K_1 -凸的, 我们有

$$\lambda z^1 + (1 - \lambda)z^2 \in G(\lambda y^1 + (1 - \lambda)y^2) + K_1. \quad (6.4.10)$$

另外, 由 F 在 S 上是凸的, 则

$$\lambda y^1 + (1 - \lambda)y^2 \in F(\lambda x^1 + (1 - \lambda)x^2).$$

根据(6.4.10)和上式,有

$$\lambda z^1 + (1 - \lambda)z^2 \in G(F(\lambda x^1 + (1 - \lambda)x^2) + K_1).$$

由此得知 θ 在 S 上是 R_+ -凸的, $G \circ F$ 在 S 上是 K_1 -凸的. 因为

$$x^* \in \partial \theta(x, \langle y^*, y \rangle)_1,$$

由定理 6.2.12 可得

$$\begin{aligned} \langle y^*, y' \rangle - \langle y^*, y \rangle &\geq \langle x^*, x' - x \rangle \\ \forall x' \in S, \forall y' \in F(x'). \end{aligned} \quad (6.4.11)$$

再根据 $y^* \in \partial_w G(y, z)_{p_1}$, 由定理 6.2.12 可知, 存在 $q_1^* \in K_1^* \setminus \{0\}$ 使得

$$\begin{aligned} \langle q_1^*, z' - z \rangle &\geq \langle y^*, y' - y \rangle \langle q_1^*, p_1 \rangle \\ \forall y' \in Y, \forall z' \in G(y'). \end{aligned}$$

从上式和(6.4.11),我们有

$$\begin{aligned} \langle q_1^*, z' - z \rangle &\geq \langle x^*, x' - x \rangle \langle q_1^*, p_1 \rangle \\ \forall x' \in S, \forall z' \in G \circ F(x'). \end{aligned}$$

根据定理 6.2.11 的(2), 便得 $x^* \in \partial_w G \circ F(x, z)_{p_1}$.

现设 $x^* \in \partial_w G \circ F(x, z)_{p_1}$, 作集合

$$S_1 = \{(y', z') \in \mathcal{Y} \times \mathcal{Z} \mid y' \in Y, z' = \bar{z} + q_1' - z,$$

$$\bar{z} \in G(y'), q_1' \in K_1\},$$

$$S_2 = \{(y', z') \in \mathcal{Y} \times \mathcal{Z} \mid x' \in S, y' \in F(x'),$$

$$z' = \langle x^*, x' - x \rangle p_1\}.$$

由于 S 和 Y 是凸的, F 在 S 上是凸的, G 在 Y 上是 K_1 -凸的, 容易

推知 S_1 和 S_2 是 $\mathscr{Y} \times \mathscr{Z}$ 中的凸集. 由已知 G 在 y 处是 K_1 -有界逆开的, 根据引理 6.2.1 得 $\text{int}S_1 \neq \emptyset$. 下面证明

$$S_2 \cap \text{int}S_1 = \emptyset. \quad (6.4.12)$$

事实上, 否则, 假若存在 $\bar{x} \in S$, $\bar{y} \in F(\bar{x})$, $\bar{z} \in G(\bar{y})$ 和 $q'_1 \in K_1$, 使得

$$y' = \bar{y}, z' = \langle x^*, \bar{x} - x \rangle p_1 = \bar{z} + q'_1 - z, \quad (6.4.13)$$

并且 $(y', z') \in \text{int}S_1$, 则存在正数 $\lambda > 0$ 使得 $(y', z' - \lambda p_1) \in S_1$, 故又存在 $\bar{z}' \in G(y')$ 和 $\bar{q}'_1 \in K_1$, 使得

$$z' - \lambda p_1 = \bar{z}' + \bar{q}'_1 - z.$$

由 (6.4.13) 和上式可知 $\bar{z}' \in G \circ F(\bar{x})$ 和

$$\bar{z}' + \bar{q}'_1 - z + \lambda p_1 = \langle x^*, \bar{x} - x \rangle p_1.$$

整理上式, 得

$$z - \langle x^*, x \rangle p_1 - \{\bar{z}' - \langle x^*, \bar{x} \rangle p_1\} = \lambda p_1 + \bar{q}'_1 \in \text{int}K_1.$$

由此, 按定义 6.1.3 的 (2) 得到 $x^* \notin \partial_w G \circ F(x, z)_{p_1}$, 这与已设矛盾. 由 (6.4.12), 根据定理 1.3.7 的 (1) 可知, 存在 $(y^*, z^*) \in \mathscr{Y}^* \times \mathscr{Z}^*$ 和 $(y^*, z^*) \neq \theta$, 使得

$$\begin{aligned} \langle y^*, y' \rangle + \langle z^*, z' + q_1 - z \rangle &\geq \langle y^*, \bar{y} \rangle + \langle z^*, \langle x^*, \bar{x} - x \rangle p_1 \rangle \\ \forall y' \in Y, \forall z' \in G(y'), \forall q_1 \in K_1, \forall \bar{x} \in S, \forall \bar{y} \in F(\bar{x}). \end{aligned} \quad (6.4.14)$$

在 (6.4.14) 中令 $z' = z$, $\bar{x} = x$, $y' = y$, $\bar{y} = y$, 得到

$$\langle z^*, q_1 \rangle \geq 0 \quad \forall q_1 \in K_1.$$

易知有 $z^* \neq \theta$ (因为若 $z^* = \theta$, 则从 $(y^*, z^*) \neq \theta$ 得 $y^* \neq \theta$. 由 (6.4.14) 可知

$$\langle y^*, y' \rangle \geq \langle y^*, \bar{y} \rangle \quad \forall y' \in Y, \forall \bar{y} \in F(\bar{x}),$$

故得对任意的 $y' \in Y$ 有 $\langle y^*, y' - y \rangle \geq 0$. 类似于定理 6.2.4 的证明, 导致与 $y \in \text{int} Y$ 相矛盾. 于是, 得知有 $z^* \in K_1^* \setminus \{0\}$. 对任意的 $y' \in Y$ 和任意的 $z' \in G(y')$ 在 (6.4.14) 中令 $\bar{y} = y, \bar{x} = x, q_1 = 0$, 则有

$$\langle z^*, z' - z \rangle \geq \langle -y^*, y' - y \rangle \quad \forall y' \in Y, \forall z' \in G(y').$$

令 $\bar{y}^* = -y^* / \langle q_1^*, p_1 \rangle, q_1^* = z^*$, 注意 $\langle q_1^*, p_1 \rangle > 0$, 上式即为

$$\langle q_1^*, z' - z \rangle \geq \langle \bar{y}^*, y' - y \rangle \langle q_1^*, p_1 \rangle$$

$$\forall y' \in Y, \forall z' \in G(y').$$

由此, 根据定理 6.2.11 的 (2) 得 $\bar{y}^* \in \partial G(y, z)_{p_1}$. 对任意的 $\bar{x} \in S$ 和任意的 $\bar{y} \in F(\bar{x})$ 在 (6.3.14) 中令 $q_1 = 0, y' = y, z' = z$, 可得

$$\langle \bar{y}^*, \bar{y} - y \rangle \geq \langle x^*, \bar{x} - x \rangle \quad \forall \bar{x} \in S, \forall \bar{y} \in F(\bar{x}).$$

于是, 从定理 6.2.11 的 (2) 便得 $x^* \in \partial \theta(x, \langle \bar{y}^*, y \rangle)_1$, 由 \bar{y}^* 的任意性, 故知 (6.4.9) 成立. \square

若 $\mathcal{X} = R, \mathcal{Y} = R^n, F: S \rightarrow R^n$ 是单值可微向量函数, 则 (6.4.9) 成为

$$\partial_w G \circ F(x, z)_{p_1} = \{y^* \cdot JF(x) \mid y^* \in \partial_w G(y, z)_{p_1}\}.$$

定义 6.4.2 设 $Y \subset \mathcal{Y}$ 是 K -凸集, $G: Y \rightarrow 2^Z, y \mapsto G(y)$ 是集值映射. 若对任意的 $y^1, y^2 \in Y$, 任意的 $\lambda \in (0, 1)$, 存在 $y_\lambda \in Y, k \in K_1$, 使得 $\lambda y^1 + (1 - \lambda)y^2 = y_\lambda + k$ 有

$$\lambda G(y^1) + (1 - \lambda)G(y^2) \subset G(y_\lambda) + K_1,$$

则称集值映射 G 在 Y 上关于 K 是 K_1 -凸的.

显然, 若 Y 是凸集, G 在 Y 上关于 K 是 K_1 -凸的, 则 G 在 Y 上也是 K_1 -凸的. 反过来, 若 Y 是凸集, G 在 Y 上是 K_1 -凸的, 则 G 在 Y 上关于 K 也是 K_1 -凸的.

定理 6.4.5 设 $S \subset \mathcal{X}$ 是非空凸集, 集合 $F(x) \subset \mathcal{Y} (\forall x \in S)$ 是非空的, $Y = F(S) = \bigcup_{x \in S} F(x)$, 集合 $G(y) \subset \mathcal{Z} (\forall y \in Y)$ 是

非空的, $x \in S$, $y \in F(x)$, $z \in G(y)$, $y \in \text{int} Y$, $p_1 \in \text{int} K_1$. 又设集值映射 $F: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto F(x)$ 在 S 上是 K -凸的, 集值映射 $G: Y \rightarrow 2^Z$, $y \mapsto G(y)$ 在 Y 上关于 K 是 K_1 -凸的. 对于集值映射

$$G \circ F: S \rightarrow 2^Z, x \mapsto G \circ F(x) = G(F(x)) = \bigcup_{y \in F(x)} G(y),$$

若 G 在 y 处是 K -有界逆开的或连通的, 则

$$\partial_{\omega} G \circ F(x, z)_{p_1} = \{\partial \theta_1(x, \langle y^*, y \rangle)_1 \mid y^* \in \partial_{\omega} G_1(y, z)_{p_1}\},$$

其中集值函数 $\theta_1: S \rightarrow 2^R$, $x \mapsto \theta_1(x) = \{\langle y^*, y \rangle \mid y \in F(x) + K\}$, 集值映射 $G_1: Y + K \rightarrow 2^Z$, $y \mapsto G_1(y)$. 若存在 $y_1 \in Y$, $q_1 \in K$ 使得 $y = y_1 + q_1$, 则 $G_1(y) = G(y_1)$.

证明 作集值映射 $F_1: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto F_1(x)$, $F_1(x) = F(x) + K$. 由于 F 在 S 上是 K -凸的, 则 F_1 在 S 上是凸的. 由 G_1 的定义可得:

$$\begin{aligned} G \circ F(x) &= \bigcup_{y \in F(x)} G(y) = \bigcup_{y \in F(x) + K} G_1(y) \\ &= \bigcup_{y \in F_1(x)} G_1(y) = G_1 \circ F_1(x). \end{aligned}$$

因为 G 在 Y 上关于 K 是 K_1 -凸的, 按定义 6.4.2 可以推得 G_1 在 $Y + K$ 上是 K_1 -凸的, 且由 F 在 S 上是 K_1 -凸的知 $Y + K$ 是凸集. 由 $y \in \text{int} Y$ 有 $y \in \text{int}(Y + K)$. 显然, G_1 在 y 处是 K_1 -有界逆开的或连通的. 由定理 6.4.4 可知结论成立. \square

§ 6.5 锥方向导数和锥弱方向导数

我们知道, 设 $S \subset \mathcal{X}$ 是非空凸集, 集值映射 $\psi: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \psi(x)$ 在 S 上是 K -凸的, 则 $K\text{-epi}\psi$ 是凸集. 又设点 $x^0 \in S$, $y^0 \in \psi(x^0)$, α 是实数, 则

$$T_{K\text{-epi}\psi}(x^0, y^0) = \text{cl} \left\{ \bigcup_{\alpha > 0} \alpha [K\text{-epi}\psi - (x^0, y^0)] \right\} \quad (6.5.1)$$

是 K -epi ψ 在点 (x^0, y^0) 处的切锥. (6.5.1) 等价于: $(d, r) \in T_{K\text{-epi}\psi}(x^0, y^0)$ 当且仅当存在序列 $\{d^k\} \subset \mathcal{X}$, $\{r^k\} \subset \mathcal{Y}$ 和 $\{\lambda_k\} \subset (0, +\infty)$, 当 $k \rightarrow \infty$ 时有 $d^k \rightarrow d$, $r^k \rightarrow r$, $\lambda_k \rightarrow 0$ 使得 $y^0 + \lambda_k r^k \in \psi(x^0 + \lambda_k d^k) + K$ ($k = 1, 2, \dots$). 由此, 从 (6.5.1) 得知

$$K\text{-epi}\psi - (x^0, y^0) \subset T_{K\text{-epi}\psi}(x^0, y^0).$$

定义 6.5.1 设集合 $S \subset \mathcal{X}$ 非空, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, 点 $x^0 \in S$, $y^0 \in \psi(x^0)$, $d \in \mathcal{X}$. 记集合

$$\Psi(x^0, y^0, d) = \{r \in \mathcal{Y} \mid (d, r) \in T_{K\text{-epi}\psi}(x^0, y^0)\}. \quad (6.5.2)$$

(1) 若 $y'(x^0, y^0; d) \in \Psi(x^0, y^0, d)$, 并且不存在 $y \in \Psi(x^0, y^0, d)$ 使得

$$y'(x^0, y^0; d) - y \in K \setminus \{0\},$$

则称 $y'(x^0, y^0; d)$ 是集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 K -方向导数. ψ 在 (x^0, y^0) 处沿 d 的所有 K -方向导数组成的集合记作 $\psi'(x^0, y^0; d)$.

(2) 若 $y'(x^0, y^0; d) \in \Psi(x^0, y^0, d)$, 并且不存在 $y \in \Psi(x^0, y^0, d)$ 使得

$$y'(x^0, y^0; d) - y \in \text{int}K,$$

则称 $y'(x^0, y^0; d)$ 是集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 K -弱方向导数. ψ 在 (x^0, y^0) 处沿方向 d 的所有 K -弱方向导数组成的集合记作 $\psi'_w(x^0, y^0; d)$.

显然有 $\psi'(x^0, y^0; d) \subset \psi'_w(x^0, y^0; d)$. 当 $\mathcal{Y} = R$, $K = R_+$ 时, 则 $\psi'(x^0, y^0; d) = \psi'_w(x^0, y^0; d)$.

注 6.5.1 特别地, 若 $\mathcal{X} = \mathcal{B}$ 是 Banach 空间, $\psi = f: S \rightarrow R \cup \{+\infty\}$ 是实值凸函数, 并且 f 在 $x^0 \in S$ 处满足 Lipschitz 条件, 则由定义 5.2.2、定义 5.3.1 和定义 5.4.1 可知, 这里的锥方向导数 $f'(x^0, f(x^0); d)$ 与 D -切导数或 C -切导数是一致的. 并且, 由

定理 5.4.2 还可知, 锥方向导数 $f'(x^0, f(x^0); d)$ 等于第 4 章定义的 G -方向导数, 即 $f'(x^0, f(x^0); d) = f^0(x^0; d)$.

以下讨论锥方向导数和锥弱方向导数与锥次微分和锥弱次微分的关系.

定义 6.5.2 设 $Y \subset \mathscr{X}$ 是非空集合, $K \subset \mathscr{X}$ 是闭凸锥.

(1) 若存在点 $a \in \mathscr{X}$ 使得 $Y \subset a + K$, 则称 Y 是 K -有界集, 或 Y 是 K -有界的.

(2) 若 $Y + K$ 是闭集, 则称 Y 是 K -闭集, 或 Y 是 K -闭的.

引理 6.5.1 设集合 $Y \subset \mathscr{X}$ 非空. 若 Y 是 K -闭的和 K -有界的, 则 Y 是 K -半紧的.

证明 见文献[27]的命题 3.3.5. \square

引理 6.5.2 设集合 $Y \subset \mathscr{X}$ 非空, $K \subset \mathscr{X}$ 是内部非空的尖闭凸锥, Y 是 K -半紧的.

(1) $Y \subset \mathscr{E}(Y, K) + K$.

(2) $Y \subset \mathscr{E}_w(Y, K) + \text{int}K \cup \{0\}$.

证明 见文献[27]的定义 3.5.1 和定理 3.5.4. \square

设 $x \in S$, $y \in \phi(x)$, 记

$$\hat{D} = \{d \in \mathscr{X} \mid \Psi(x, y, d) \neq \emptyset\}.$$

若 $S = \mathscr{X}$, 则显然有 $\hat{D} = \mathscr{X}$.

定理 6.5.3 设 $S \subset \mathscr{X}$ 是非空凸集, $\phi: S \rightarrow 2^{\mathscr{X}}$ 是集值映射, $x \in S$, $y \in \phi(x)$, $\hat{D} = S$, $p \in \text{int}K$. 若 ϕ 在 S 上是 K -凸的, 并且 $\Psi(x, y, 0)$ 是 K -有界的, 则对任意的 $d \in \mathscr{X}$ 有 $\psi'(x, y; d) \neq \emptyset$ 和 $\psi'_w(x, y; d) \neq \emptyset$, 并且

$$\partial\psi(x, y)_p$$

$$\supset \left\{ x^* \in \mathscr{X}^* \mid 0 \in \mathscr{E} \left(\bigcup_{d \in \mathscr{X}} [\psi'(x, y; d) - \langle x^*, d \rangle p], K \right) \right\},$$

$$\partial_w\psi(x, y)_p$$

$$= \left\{ x^* \in \mathscr{X}^* \mid 0 \in \mathscr{E}_w \left(\bigcup_{d \in \mathscr{X}} [\psi'_w(x, y; d) - \langle x^*, d \rangle p], K \right) \right\}.$$

证明 取任意的 $d^1, d^2 \in \mathcal{D}$, 任意的 $r^1, r^2 \in \Psi(x, y, d)$ 由 (6.5.2) 知 $(d^1, r^1), (d^2, r^2) \in T_{K\text{-epi}\psi}(x, y)$. 因为 $T_{K\text{-epi}\psi}(x, y)$ 是凸锥, 故 $(d^1 + d^2, r^1 + r^2) \in T_{K\text{-epi}\psi}(x, y)$, 再由 (6.5.2) 有 $r^1 + r^2 \in \Psi(x, y, d^1 + d^2)$. 因此, 得到

$$\Psi(x, y, d^1) + \Psi(x, y, d^2) \subset \Psi(x, y, d^1 + d^2) \quad \forall d^1, d^2 \in \mathcal{D}.$$

对任意的 $d \in \mathcal{D}$, 令 $d^1 = d, d^2 = -d$, 取点 $r \in \Psi(x, y, d^2)$, 我们有

$$\Psi(x, y, d) + r \subset \Psi(x, y, 0).$$

因为 $\Psi(x, y, 0)$ 是 K -有界的, 所以由上式可知 $\Psi(x, y, d)$ 也是 K -有界的. 另外, 由 $T_{K\text{-epi}\psi}(x, y, d)$ 是闭的, 易知 $\Psi(x, y, d)$ 也是闭的, 又因 K 是闭的, 从而 $\Psi(x, y, d)$ 是 K -闭的. 由此, 根据引理 6.5.1 可知 $\Psi(x, y, d)$ 是 K -半紧的, 再由引理 6.2.5 得知 $\psi'(x, y; d) \neq \emptyset$ 和 $\psi'_w(x, y; d) \neq \emptyset$. 于是, 根据引理 6.5.2 得到

$$\Psi(x, y, d) \subset \psi'(x, y; d) + K, \quad (6.5.3)$$

$$\Psi(x, y, d) \subset \psi'_w(x, y; d) + K. \quad (6.5.4)$$

记集合

$$S_1 = \left\{ x^* \in \mathcal{X}^* \mid 0 \in \mathcal{E} \left(\bigcup_{(d, r) \in T_{K\text{-epi}\psi}(x, y)} [r - \langle x^*, d \rangle p], K \right) \right\}, \quad (6.5.5)$$

$$S_2 = \left\{ x^* \in \mathcal{X}^* \mid 0 \in \mathcal{E} \left(\bigcup_{d \in \mathcal{D}} [\psi'(x, y; d) - \langle x^*, d \rangle p], K \right) \right\}, \quad (6.5.6)$$

$$S_3 = \left\{ x^* \in \mathcal{X}^* \mid 0 \in \mathcal{E}_w \left(\bigcup_{(d, r) \in T_{K\text{-epi}\psi}(x, y)} [r - \langle x^*, d \rangle p], K \right) \right\}, \quad (6.5.7)$$

$$S_4 = \left\{ x^* \in \mathcal{X}^* \mid 0 \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{D}} [\psi'_w(x, y; d) - \langle x^*, d \rangle p], K \right) \right\}. \quad (6.5.8)$$

先证明 $S_1 = S_2$. 设 $x^* \in S_2$, 假若 $x^* \notin S_1$, 则存在 $(d', r') \in T_{K\text{-epi}\psi}(x, y)$, 使得

$$0 - (r' - \langle x^*, d' \rangle p) \in K \setminus \{0\}. \quad (6.5.9)$$

另外, 由 (6.5.2) 和 (6.5.3) 知 $r' \in \psi'(x, y; d') + K$, 故存在 $q' \in K$, $y' \in \psi'(x, y; d')$ 使 $r' = y' + q'$. 再由 (6.5.9) 有

$$0 - (y' - \langle x^*, d' \rangle p) \in K \setminus \{0\} + q' \subset K \setminus \{0\},$$

于是得知 $x^* \notin S_2$, 这与已设矛盾. 反之, 设 $x^* \in S_1$, 则不存在 $(d, r) \in T_{K\text{-epi}\psi}(x, y)$ 使得

$$- (r - \langle x^*, d \rangle p) \in K \setminus \{0\}. \quad (6.5.10)$$

由定义 6.5.1 知 $\psi'(x, y; d) \subset \Psi(x, y, d)$, 故再由 (6.5.10) 可知不存在 $r \in \psi'(x, y; d)$, 使 $- (r - \langle x^*, d \rangle p) \in K \setminus \{0\}$, 从而 $x^* \in S_2$.

同证明 $S_1 = S_2$ 类似, 可由 (6.5.4) 推得 $S_3 = S_4$. 按定义 6.1.2 可知 $S_1 \subset S_3$, 从而有 $S_2 \subset S_4$. 现证 $S_2 \subset \partial\psi(x, y)_p$. 设 $x^* \in S_2 = S_1$, 从 (6.5.1) 知对任意的 $x' \in S$ 和 $y' \in \psi(x')$ 有 $(x' - x, y' - y) \in T_{K\text{-epi}\psi}(x, y)$. 于是, 由 (6.5.5) 可得 $y - y' + \langle x^*, x' - x \rangle p \notin K \setminus \{0\}$, 再由定义 6.1.3 的 (1) 便知 $x^* \in \partial\psi(x, y)_p$. 类似地, 由 (6.5.7) 和 (6.5.8) 可以推得 $S_4 \subset \partial_w\psi(x, y)_p$.

最后, 证明 $\partial_w\psi(x, y)_p \subset S_4$. 用反证法, 假设 $x^* \in \partial\psi(x, y)_p$ 而 $x^* \notin S_4$, 则有 $x^* \notin S_3$. 由此, 存在 $(d', r') \in T_{K\text{-epi}\psi}(x, y)$ 使得

$$0 - (r' - \langle x^*, d' \rangle p) \in \text{int}K, \quad (6.5.11)$$

因而有数列 $\{d^k\} \subset \mathcal{H}$, $\{r^k\} \subset \mathcal{D}$ 和 $\{\lambda_k\} \subset (0, +\infty)$, 当 $k \rightarrow \infty$ 时, 有 $d^k \rightarrow d'$, $r^k \rightarrow r'$, $\lambda_k \rightarrow 0$, 使得

$$y + \lambda_k r^k \in \psi(x + \lambda_k d^k) + K, \quad k = 1, 2, \dots.$$

据此, 存在 $y_k \in \psi(x + \lambda d^k)$ 和 $p^k \in K$ 使 $y_k + p^k = y + \lambda_k r^k$ ($k = 1, 2, \dots$), 或有

$$r^k = (y_k - y + p^k)/\lambda_k, \quad k = 1, 2, \dots$$

令 $x^k = x + \lambda_k d^k$, 则 $d^k = (x^k - x)/\lambda_k$ ($k = 1, 2, \dots$), $y^k \in \psi(x^k)$. 因为当 $k \rightarrow \infty$ 时, $r^k \rightarrow r'$, $d^k \rightarrow d'$, 故对充分大的 k' , 由 (6.5.11) 得知

$$-(y^{k'} - y + p^{k'})/\lambda_{k'} + \langle x^*, (x^{k'} - x)/\lambda_{k'} \rangle p \in \text{int}K,$$

从而有 $y - y^{k'} + \langle x^*, x^{k'} - x \rangle p \in \text{int}K$. 于是, 根据定义 6.1.3 的 (2) 得 $x^* \notin \partial\psi(x, y)_r$, 导致与已设矛盾. \square

定理 6.5.4 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S$, $y \in \psi(x)$, $\hat{D} = S$, $p \in \text{int}K$. 若 ψ 在 S 上是 K -凸的, 并且 $\Psi(x, y, \theta)$ 是 K -有界的, 则对任意的 $d \in \mathcal{X}$, 有 $\psi'(x, y; d)$ 和 $\psi'_w(x, y; d)$ 关于 d 在 \mathcal{X} 上是 K -凸的.

证明 由定理 6.5.3 的证明可得

$$\Psi(x, y, d^1) + \Psi(x, y, d^2) \subset \Psi(x, y, d^1 + d^2)$$

$$\forall d^1, d^2 \in \mathcal{X}.$$

再由 (6.5.3) 和 (6.5.4) 得到

$$\psi'(x, y; d^1) + \psi'(x, y; d^2) \subset \psi'(x, y; d^1 + d^2) + K,$$

$$\psi'_w(x, y; d^1) + \psi'_w(x, y; d^2) \subset \psi'_w(x, y; d^1 + d^2) + K.$$

因此, 定理的结论成立. \square

定理 6.5.5 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S$, $y \in \psi(x)$, $\hat{D} = S$, $p \in \text{int}K$. 又设 ψ 在 S 上是 K -凸的, 并且 $\Psi(x, y, \theta)$ 是 K -有界的.

(1) 若存在 $q^* \in \text{int}K^*$, 使得

$$\langle q^*, r \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall (d, r) \in T_{K\text{-epi}\psi}(x, y),$$

则 $x^* \in \partial\psi(x, y)_r$.

(2) 若存在 $q^* \in \text{int}K^*$, 使得

$$\langle q^*, y' \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle$$

$$\forall d \in \mathcal{K}, \forall y' \in \psi(x, y; d),$$

则 $x^* \in \partial \psi(x, y)_p$.

(3) 若存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, r \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall (d, r) \in T_{K\text{-epi}\psi}(x, y),$$

则 $x^* \in \partial_w \psi(x, y)_p$.

(4) 若存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle$$

$$\forall d \in \mathcal{K}, \forall y' \in \psi_w(x, y; d),$$

则 $x^* \in \partial_w \psi(x, y)_p$.

证明 (1) 因为对任意的 $x' \in S, y' \in \psi(x')$, 有 $(x' - x, y' - y) \in T_{K\text{-epi}\psi}(x, y)$, 所以根据定理 6.2.11 的(1)可得结论.

(2) 反证: 假若 $x^* \notin \partial \psi(x, y)_p$, 由定理 6.5.3 可知, 存在 $\bar{y} \in \psi(x, y; d)$ 使得

$$0 - (\bar{y} - \langle x^*, d \rangle p) \in K \setminus \{0\}.$$

再由 $q^* \in \text{int} K^*$, 得 $\langle q^*, \bar{y} \rangle < \langle x^*, d \rangle \langle q^*, p \rangle$, 与已知矛盾.

(3) 与(1)的证明类似.

(4) 与(2)的证明类似. \square

定理 6.5.6 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S, y \in \psi(x), \hat{D} = S, p \in \text{int} K$. 又设 ψ 在 S 上是 K -凸的, 并且 $\Psi(x, y, \theta)$ 是 K -有界的. 若 $x^* \in \partial_w \psi(x, y)_p$, 则存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle$$

$$\forall d \in \mathcal{K}, \forall y' \in \psi(x, y; d),$$

$$\langle q^*, y' \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle$$

$$\forall d \in \mathcal{X}, \forall y' \in \psi'_w(x, y; d),$$

$$\langle q^*, r \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall (d, r) \in T_{K\text{-epi}\psi}(x, y).$$

证明 只需证明最后一个式子成立. 设 $x^* \in \partial_w \psi(x, y)_p$, 由定理 6.5.4 的证明可知 $x^* \in S_3$. 作集合

$$E = \{y' \in \mathcal{Y} \mid y' = -r + \langle x^*, d \rangle, (d, r) \in T_{K\text{-epi}\psi}(x, y)\},$$

则 E 是凸集, 并且 $E \cap \text{int}K = \emptyset$. 由此, 利用定理 1.3.7 的(1)可知, 存在 $q^* \neq 0$ 使得

$$\langle q^*, q \rangle \geq \langle q^*, y' \rangle \quad \forall q \in K, \forall y' \in E.$$

由于 $0 \in E \cap K$, 故对任意的 $q \in K$, 有 $\langle q^*, q \rangle \geq 0$, 因此 $q^* \in K^* \setminus \{0\}$. 再由 $0 \geq \langle q^*, y' \rangle (y' \in E)$, 即得结论. \square

定理 6.5.7 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S, y \in \psi(x), \hat{D} = S, p \in \text{int}K$. 又设 ψ 在 S 上是 K -凸的, 并且 $\Psi(x, y, 0)$ 是 K -有界的. 若存在 $\bar{x}^* \in \partial_w \psi(x, y)_p$ 和 $\bar{y} \in \psi'(x, y; \bar{d})$, 使得 $\bar{y} = \langle \bar{x}^*, \bar{d} \rangle p$, 则

$$\bar{y} = \left(\max \{ \langle x^*, \bar{d} \rangle \mid x^* \in \partial_w \psi(x, y)_p \} \right) p. \quad (6.5.12)$$

证明 作集合

$$E = \{y' \in \mathcal{Y} \mid y' = -\bar{y} + \langle x^*, \bar{d} \rangle p, x^* \in \partial_w \psi(x, y)_p\},$$

注意: \bar{y} 和 \bar{d} 是给定的. 从定理 6.3.2 知 $\partial_w \psi(x, y)_p$ 是凸的, 故 E 是凸的. 由定理 6.5.3 可推得 $E \cap \text{int}K = \emptyset$, 再利用定理 1.3.7, 即知存在 $q^* \neq 0$ 使得

$$\langle q^*, q \rangle \geq \langle q^*, y' \rangle \quad \forall q \in K, \forall y' \in E.$$

由于 $0 \in E \cap K$, 故对任何 $q \in K$ 有 $\langle q^*, q \rangle \geq 0$, 因此 $q^* \in K^* \setminus \{0\}$. 上式对任何 $y' = -\bar{y} + \langle x^*, \bar{d} \rangle p \in E$, 有

$$\langle q^*, \bar{y} \rangle \geq \langle x^*, \bar{d} \rangle \langle q^*, p \rangle \quad \forall x^* \in \partial_w \psi(x, y)_p.$$

从已知 $\bar{y} = \langle \bar{x}^*, \bar{d} \rangle_p$ 代入上式, 得

$$\langle \bar{x}^*, \bar{d} \rangle \geq \langle x^*, \bar{d} \rangle \quad \forall x^* \in \partial_w \psi(x, y)_p,$$

因而 (6.5.12) 成立. \square

定义 6.5.3 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, 点 $x^0 \in S, y^0 \in \psi(x^0)$. 若 ψ 在 S 上是 K -凸的, 则称集合

$$\begin{aligned} N_{K\text{-epi}\psi}(x^0, y^0) \\ = \{ (x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^* \mid \langle (x^*, y^*), (d, r) \rangle \leq 0, \\ \forall (d, r) \in T_{K\text{-epi}\psi}(x^0, y^0) \} \end{aligned}$$

是集值映射 ψ 在点 (x^0, y^0) 处的法锥.

定理 6.5.8 设 $S \subset \mathcal{X}$ 是非空凸集, $\psi: S \rightarrow 2^{\mathcal{Y}}$ 是集值映射, $x \in S, y \in \psi(x), \hat{D} = S, p \in \text{int}K$. 又设 ψ 在 S 上是 K -凸的, 并且 $\Psi(x, y, \theta)$ 是 K -有界的.

(1) $x^* \in \partial_w \psi(x, y)_p$ 当且仅当存在 $q^* \in K^* \setminus \{0\}$, 使得

$$(\langle q^*, p \rangle x^*, -q^*) \in N_{K\text{-epi}\psi}(x, y).$$

(2) 若存在 $q^* \in \text{int}K^*$, 使得

$$(\langle q^*, p \rangle x^*, -q^*) \in N_{K\text{-epi}\psi}(x, y),$$

则 $x^* \in \partial \psi(x, y)_p$.

证明 由定义 6.5.3、定理 6.5.5 和定理 6.5.6 可得证. \square

第 7 章 一般集值映射的 锥弱次微分

在上一章,我们研究了锥凸集值映射的锥次微分问题.本章将对一般的集值映射引进两种锥弱次微分概念,并讨论它们的一些相应的性质.

§ 7.1 有效 Hahn-Banach 定理

在第 4 章,我们看到了 Hahn-Banach 定理在证明 Lipschitz 函数的次微分的存在性以及其它性质中起到了关键的作用.现在先证明关于集值映射的有效 Hahn-Banach 定理,为后面研究一般集值映射的锥弱次微分打下基础.

设 \mathcal{X} 和 \mathcal{Y} 是局部凸线性拓扑空间, \mathcal{X}^* 和 \mathcal{Y}^* 分别是 \mathcal{X} 和 \mathcal{Y} 的对偶空间, 集合 $S \subset \mathcal{X}$, $\psi: S \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \psi(x)$ 是集值映射. 又设 $K \subset \mathcal{Y}$ 是内部非空的尖闭凸锥, $p \in \text{int}K$, K^* 是 K 的对偶锥.

定义 7.1.1 设集合 $S \subset \mathcal{X}$ 非空, 点 $x^0 \in \text{int}S$. 若存在一点 a 和 x^0 的一个邻域 $U(x^0)$, 使得 $\psi(x) \subset a + K$ ($\forall x \in U(x^0)$), 则称集值映射 ψ 在点 x^0 附近是局部 K -有界的.

引理 7.1.1 设集合 $S \subset \mathcal{X}$ 非空, $x^0 \in \text{int}S$. 若 ψ 在 x^0 附近是局部 $(-K)$ -有界的, 则 $\text{int}(K\text{-epi}\psi) \neq \emptyset$, 其中 $K\text{-epi}\psi$ 是 ψ 在 S 上的 K -上图象.

证明 与引理 6.2.1 的证明类似. \square

定理 7.1.2 设 $\mathcal{L} \subset \mathcal{X}$ 是非空线性子空间, ψ 在 \mathcal{L} 上是

K -凸的, ϕ 在 $x^0 \in \mathcal{X}$ 附近是局部 $(-K)$ -有界的 (或 K -有界逆开, 或连通的). 若存在点 $x_0^* \in \mathcal{X}^*$, 使得

$$0 \in \mathcal{E}_w \left(\bigcup_{x \in \mathcal{L}} [\phi(x) - \langle x_0^*, x \rangle p], K \right),$$

则存在线性泛函 $\bar{x}^* \in \mathcal{X}^*$ 使得

$$\langle x_0^*, x \rangle = \langle \bar{x}^*, x \rangle \quad \forall x \in \mathcal{L},$$

并且

$$0 \in \mathcal{E}_w \left(\bigcup_{x \in \mathcal{L}} [\phi(x) - \langle \bar{x}^*, x \rangle p], K \right).$$

证明 设集合

$$S_1 = \{(x, y) | x \in \mathcal{X}, y \in \phi(x) + K\},$$

$$S_2 = \{(x, y) | x \in \mathcal{L}, y = \langle x_0^*, x \rangle p\},$$

则 S_1 和 S_2 是 $\mathcal{X} \times \mathcal{Y}$ 中的子集, 并且易知 $S_2 \neq \emptyset$. 另外, 由引理 7.2.1 (或引理 6.2.1 或引理 6.2.3) 可得 $\text{int} S_1 \neq \emptyset$. 现在证明

$$S_2 \cap \text{int} S_1 = \emptyset. \quad (7.1.1)$$

事实上, 假若有 $(x', y') \in S_2 \cap \text{int} S_1$, 因为 $p \neq 0$, 则存在 $\lambda > 0$ 使得 $(x', y' - \lambda p) \in S_1$. 因此, 有 $x' \in \mathcal{L}$, $y'' \in \phi(x')$ 和 $q' \in K$, 使得

$$\langle x_0^*, x' \rangle p - \lambda p = y'' + q'.$$

由于 $p \in \text{int} K$, 故有

$$0 - (y'' - \langle x_0^*, x' \rangle p) = \lambda p + q' \in \text{int} K.$$

据此, 由定义 6.1.2 的 (2) 得

$$0 \notin \mathcal{E}_w \left(\bigcup_{x \in \mathcal{L}} [\phi(x) - \langle x_0^*, x \rangle p], K \right),$$

这导致与已知矛盾. 故 (7.1.1) 成立. 现由 (7.1.1), 利用定理 1.3.7 得知, 存在 $(x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^*$ 和 $(x^*, y^*) \neq 0$ 使得

$$\langle x^*, x \rangle + \langle y^*, y \rangle + q \geq \langle x^*, x_0 \rangle + \langle y^*, \langle x_0^*, x_0 \rangle p \rangle$$

$$\forall x \in \mathcal{X}, \forall y \in \phi(x), \forall q \in K, \forall x_0 \in \mathcal{L}. \quad (7.1.2)$$

由于已知 $\theta \in \bigcup_{x \in \mathcal{L}} [\phi(x) - \langle x_0^*, x \rangle p]$, 从而存在 $\bar{x} \in \mathcal{L}$ 使得 $\bar{y} \in \phi(\bar{x})$ 和 $\bar{y} = \langle x_0^*, \bar{x} \rangle p$, 因此 $(\bar{x}, \bar{y}) \in S_2 \cap S_1$. 由(7.1.2)可得

$$\langle y^*, q \rangle \geq 0 \quad \forall q \in K.$$

因为 \mathcal{L} 是线性子空间, $y^* \neq \theta$, 所以 $y^* \in K^* \setminus \{\theta\}$. 再由 $p \in \text{int}K$ 和定理 1.4.11 的(1)可得 $\langle y^*, p \rangle > 0$. 又由于 \mathcal{L} 是线性子空间, 故由(7.1.2)得到

$$\langle x^*, x_0 \rangle + \langle x_0^*, x_0 \rangle \langle y^*, p \rangle = 0 \quad \forall x_0 \in \mathcal{L}.$$

设 $\bar{x}^* = -x^* / \langle y^*, p \rangle$, 从上式有

$$\langle \bar{x}^*, x_0 \rangle = \langle x_0^*, x_0 \rangle \quad \forall x_0 \in \mathcal{L},$$

再由(7.1.2)可得

$$\langle y^*, y \rangle - \langle \bar{x}^*, x \rangle \langle y^*, p \rangle \geq 0 \quad \forall x \in \mathcal{X}, \forall y \in \phi(x). \quad (7.1.3)$$

现假设

$$\theta \notin \mathcal{E}_w \left(\bigcup_{x \in \mathcal{X}} [\phi(x) - \langle \bar{x}^*, x \rangle p], K \right),$$

根据定义 6.1.2 可知存在 $\bar{x}^0 \in \mathcal{X}$, $\bar{y}^0 \in \phi(\bar{x}^0)$, 使得

$$\theta - (\bar{y}^0 - \langle \bar{x}^*, \bar{x}^0 \rangle p) \in \text{int}K.$$

因为 $y^* \in K^* \setminus \{\theta\}$, 所以得到

$$\langle y^*, -\bar{y}^0 + \langle \bar{x}^*, \bar{x}^0 \rangle p \rangle > 0,$$

这与(7.1.3)矛盾. 于是定理得证. \square

推论 7.1.3 设 $\mathcal{L} \subset \mathcal{X}$ 是非空线性子空间, ϕ 在 \mathcal{X} 上是 K -凸的, ϕ 在 $x^0 \in \mathcal{X}$ 附近是局部 $(-K)$ -有界的(或 K -有界逆

开,或连通的).若存在点 $x_0^* \in \mathcal{X}^*$ 和 $q_0^* \in K^* \setminus \{0\}$, 使得

$$\langle q_0^*, y \rangle \geq \langle x_0^*, x \rangle \langle q_0^*, p \rangle \quad \forall x \in \mathcal{L}, \forall y \in \psi(x),$$

以及 $\bar{y} = \langle x_0^*, \bar{x} \rangle p$, 其中 $\bar{y} \in \psi(\bar{x})$, 则存在线性泛函 $\bar{x}^* \in \mathcal{X}^*$ 和 $\bar{q}^* \in K^* \setminus \{0\}$, 使得

$$\langle \bar{x}^*, x \rangle = \langle x_0^*, x \rangle \quad \forall x \in \mathcal{L}$$

和

$$\langle \bar{q}^*, y \rangle \geq \langle \bar{x}^*, x \rangle \langle \bar{q}^*, p \rangle \quad \forall x \in \mathcal{X}, \forall y \in \psi(x).$$

证明 先证明有 $0 \in \mathcal{E}_w\left(\bigcup_{x \in \mathcal{L}} [\psi(x) - \langle x_0^*, x \rangle p], K\right)$. 因为否则, 由定义 6.1.2 的 (2) 可知, 存在 $x' \in \mathcal{L}$, $y' \in \psi(x')$, 使得

$$0 - (y' - \langle x_0^*, x' \rangle p) \in \text{int}K.$$

据此, 由 $p \in \text{int}K$ 和 $q_0^* \in K^* \setminus \{0\}$, 我们有

$$\langle q_0^*, -y' \rangle + \langle x_0^*, x' \rangle p > 0,$$

这导致与已知矛盾. 由已证的结果, 根据定理 7.1.2 的证明和 (7.1.3) 可知, 存在线性泛函 $\bar{x}^* \in \mathcal{X}^*$ 和 $\bar{q}^* \in K^* \setminus \{0\}$, 使得本定理的结论成立. \square

推论 7.1.4 设 $\mathcal{L} \subset \mathcal{X}$ 是非空线性子空间, $f: \mathcal{X} \rightarrow R$ 是凸函数, 并且是连续的. 若存在 $x_0^* \in \mathcal{X}^*$ 使得

$$\langle x_0^*, x \rangle \leq f(x) \quad \forall x \in \mathcal{L},$$

又 $f(\bar{x}) = \langle x_0^*, \bar{x} \rangle$, 其中 $\bar{x} \in \mathcal{X}$, 则存在线性泛函 $\bar{x}^* \in \mathcal{X}^*$ 使得

$$\langle x_0^*, x \rangle = \langle \bar{x}^*, x \rangle \quad \forall x \in \mathcal{L}$$

和

$$\langle \bar{x}^*, x \rangle \leq f(x) \quad \forall x \in \mathcal{X}.$$

证明 由推论 7.1.3 可得证. \square

为了证明集值映射在一点处的延拓成立, 以下引进集值映射的锥次线性性的概念.

定义 7.1.2 设集合 $S \subset \mathcal{X}$ 非空, $\psi: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$ 是集值映射.

(1) 对任意的 $x \in S$ 和 $\alpha > 0$, 有 $\psi(\alpha x) = \alpha\psi(x)$ (正齐次性);

(2) 对任意的 $x^1, x^2 \in S$, 有 $\psi(x^1) + \psi(x^2) \subset \psi(x^1 + x^2) + K$ (次可加性).

则称集值映射 ψ 在 S 上是 K -次线性的.

显然, ψ 在 S 上是 K -次线性的, 则 ψ 在 S 上是 K -凸的.

定理 7.1.5 设集值映射 ψ 在 \mathcal{X} 上是 K -次线性的, ψ 在 \mathcal{X} 中的某一点处附近是局部 $(-K)$ -有界的 (或 K -有界逆开的, 或连通的), $x^0 \in \mathcal{X}$, $\theta \in \psi(\theta) \subset K$. 若有 $q_0^* \in K^* \setminus \{\theta\}$ 使 $\inf_{y \in \psi(x^0)} \langle q_0^*, y \rangle > -\infty$, 则存在线性泛函 $x^* \in \mathcal{X}^*$, 使得

$$\inf \langle q_0^*, \psi(x^0) \rangle = \langle x^*, x^0 \rangle \langle q_0^*, p \rangle \quad (7.1.4)$$

和

$$\theta \in \mathcal{C}_w \left(\bigcup_{x \in \mathcal{X}} [\psi(x) - \langle x^*, x \rangle p], K \right).$$

证明 先考虑 $x^0 \neq \theta$ 的情况. 从 $p \in \text{int}K$ 和 $q_0^* \in K^* \setminus \{\theta\}$, 由定理 1.4.11 的(1)得 $\langle q_0^*, p \rangle > 0$.

设 $\alpha = \inf \langle q_0^*, \psi(x^0) \rangle / \langle q_0^*, p \rangle$, 则有

$$\langle q_0^*, y \rangle \geq \alpha \langle q_0^*, p \rangle \quad \forall y \in \psi(x^0). \quad (7.1.5)$$

因为 ψ 在 \mathcal{X} 上是 K -次线性的, 我们有

$$\psi(x^0) + \psi(-x^0) \subset \psi(\theta) + K \subset K.$$

由 $q_0^* \in K^* \setminus \{\theta\}$, 则得

$$\langle q_0^*, y \rangle + \langle q_0^*, \bar{y} \rangle \geq 0 \quad \forall y \in \psi(x^0), \forall \bar{y} \in \psi(-x^0). \quad (7.1.6)$$

下面证明有

$$\langle q_0^*, \bar{y} \rangle \geq -\alpha \langle q_0^*, p \rangle \quad \forall \bar{y} \in \psi(-x^0). \quad (7.1.7)$$

事实上, 假设(7.1.7)不成立, 则存在 $\delta > 0$ 和 $\bar{y}^0 \in \psi(-x^0)$, 使得

$$\langle q_0^*, \bar{y}^0 \rangle < -\alpha \langle q_0^*, p \rangle - \delta, \quad (7.1.8)$$

再由 α 的定义可知存在 $y^0 \in \psi(x^0)$, 使得

$$\langle q_0^*, y_0 \rangle < \alpha \langle q_0^*, p \rangle + \delta. \quad (7.1.9)$$

考虑 (7.1.8) + (7.1.9), 有

$$\langle q_0^*, y^0 \rangle + \langle q_0^*, \bar{y}^0 \rangle < 0,$$

但它与 (7.1.6) 矛盾. 另外, 由 $x^0 \neq \theta$, 故存在 $x_0^* \in \mathcal{X}^*$ 使得 $\langle x_0^*, x^0 \rangle = \alpha$, 即有

$$\inf \langle q_0^*, \psi(x^0) \rangle = \langle x_0^*, x^0 \rangle \langle q_0^*, p \rangle.$$

从 (7.1.5) 和 (7.1.7) 可以推证下面两式成立:

$$\langle x_0^*, x^0 \rangle p - y \notin \text{int}K \quad \forall y \in \psi(x^0), \quad (7.1.10)$$

$$\dots \langle x_0^*, x^0 \rangle p - y \notin \text{int}K \quad \forall y \in \psi(-x^0). \quad (7.1.11)$$

因为假若 (7.1.10) 不成立, 则存在 $y' \in \psi(x^0)$ 使

$$\langle x_0^*, x^0 \rangle p - y' \in \text{int}K,$$

从 $q_0^* \in K^* \setminus \{\theta\}$ 则得

$$\langle q_0^*, \langle x_0^*, x^0 \rangle p - y' \rangle > 0,$$

这与 (7.1.5) 矛盾. 类似可证 (7.1.11). 现设 $\mathcal{L} = \{x \in \mathcal{X} \mid x = \alpha x^0, \alpha \in R\}$, 显然 \mathcal{L} 是 \mathcal{X} 中的非空线性子空间. 我们证明有

$$\theta \in \mathcal{E}_w \left[\bigcup_{x \in \mathcal{L}} [\psi(x) - \langle x_0^*, x \rangle p], K \right]. \quad (7.1.12)$$

事实上, 假若 $\theta \notin \mathcal{E}_w \left[\bigcup_{x \in \mathcal{L}} [\psi(x) - \langle x_0^*, x \rangle p], K \right]$, 则存在 $\alpha' \in R$ 使得

$$\langle x_0^*, \alpha' x^0 \rangle p - y \in \text{int}K, \quad y \in \psi(\alpha' x^0). \quad (7.1.13)$$

下面分三种情形讨论:

(a) 当 $\alpha' = 0$ 时, 从 (7.1.13) 有 $-y \in \text{int}K$, $y \in \psi(\theta)$. 但由已知 $\psi(\theta) \subset K$ 和 K 是尖锥, 故有 $-y \notin K$, 这导致矛盾.

(b) 当 $\alpha' > 0$ 时, 从 (7.1.13) 和 $\psi(\alpha' x^0) = \alpha' \psi(x^0)$ 可得

$$\langle x_0^*, x^0 \rangle p - \frac{y}{\alpha'} \in \text{int}K, \quad \frac{y}{\alpha'} \in \psi(x^0),$$

它与 (7.1.10) 相矛盾.

(c) 当 $\alpha' < 0$ 时, 从 (7.1.13) 和 $\psi(\alpha' x^0) = -\alpha' \psi(-x^0)$ 得

$$-\langle x_0^*, x^0 \rangle p - \frac{y}{-\alpha'} \in \text{int}K, \quad \frac{y}{-\alpha'} \in \psi(-x^0),$$

它与 (7.1.11) 相矛盾.

现在考虑 $x^0 = \theta$ 的情况. 设 $\mathcal{L} = \{\theta\}$, 由于 $\theta \in \psi(\theta) \subset K$, 故 $-\psi(\theta) \notin \text{int}K$. 对 $x_0^* \in \mathcal{X}^*$ 有

$$\theta \in \mathcal{E}_w\left(\bigcup_{x \in \mathcal{L}} [\psi(x) - \langle x_0^*, x \rangle p], K\right),$$

此时显然有 (7.1.4) 成立. 根据定理 7.1.2 和 (7.1.13) 便知, 存在 $x^* \in \mathcal{X}^*$ 使得

$$\langle x^*, x^0 \rangle = \langle x_0^*, x^0 \rangle$$

和 $\theta \in \mathcal{E}_w\left(\bigcup_{x \in \mathcal{L}} [\psi(x) - \langle x_0^*, x \rangle p], K\right)$. 定理证毕. \square

由定理 7.1.5 易得下面两个推论.

推论 7.1.6 所有假设与定理 7.1.5 的相同, 则存在 $x^* \in \mathcal{X}^*$ 和 $q^* \in K^* \setminus \{\theta\}$, 使得

$$\inf \langle q_0^*, \psi(x^0) \rangle = \langle x^*, x^0 \rangle \langle q_0^*, p \rangle$$

和

$$\langle q^*, y \rangle \geq \langle x^*, x \rangle \langle q^*, p \rangle \quad \forall x \in \mathcal{X}, y \in \psi(x).$$

推论 7.1.7 设实值函数 $f: \mathcal{X} \rightarrow R$ 是连续的, 并且在 \mathcal{X} 上是 R_+ -次线性的, $p = 1$, $f(\theta) = 0$. 若存在 x^0 使得 $f(x^0) > -\infty$, 则存在线性泛函 $x^* \in \mathcal{X}^*$, 使得 $f(x^0) = \langle x^*, x^0 \rangle$ 和

$$\langle x^*, x \rangle \leq f(x) \quad \forall x \in \mathcal{X}.$$

推论 7.1.7 就是通常函数(泛函)的 Hahn-Banach 定理. 为了区别, 我们将定理 7.1.2 和定理 7.1.5 称为有效 Hahn-Banach 定理.

§ 7.2 C -锥弱次微分

本节从 C -切锥来定义集值映射的 K -上图象的 C -切锥. 由此, 引进 C -锥弱方向导数和 C -锥弱次微分.

设 \mathcal{X} 和 \mathcal{Y} 是局部凸线性拓扑空间, $x \in \mathcal{X}$, $\phi: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi(x)$ 是集值映射, 又 $K \subset \mathcal{Y}$ 是内部非空的尖闭凸锥. 由注 5.2.3 知, ϕ 在 \mathcal{X} 上的 K -上图象 $K\text{-epi } \phi$ 在点 $(x, y) \in K\text{-epi } \phi$ 处的 C -切锥 $T_{K\text{-epi } \phi}(x, y)$ 具有如下性质:

$(d, r) \in T_{K\text{-epi } \phi}^C(x, y)$ 当且仅当对任何收敛于 (x, y) 的序列 $\{(x^k, y^k)\} \subset K\text{-epi } \phi$ 和收敛于 0 的正数列 $\{t_k\}$, 存在一序列 $\{(d^k, r^k)\} \in \mathcal{X} \times \mathcal{Y}$ 收敛于 (d, r) , 使得有

$$(x^k + t_k d^k, y^k + t_k r^k) \in K\text{-epi } \phi, \quad k = 1, 2, \dots,$$

即 $y^k + t_k r^k \in \phi(x^k + t_k d^k) + K$. 由定理 5.2.1 知 $T_{K\text{-epi } \phi}^C(x, y)$ 是闭凸的.

定义 7.2.1 设集合 $\phi(x) \neq \emptyset$ ($\forall x \in \mathcal{X}$), 点 $x^0 \in \mathcal{X}$, $y^0 \in \phi(x^0)$, $d \in \mathcal{X}$. 记集合

$$\Psi(x, y, d) = \{r \in \mathcal{Y} \mid (d, r) \in T_{K\text{-epi } \phi}^C(x, y)\}. \quad (7.2.1)$$

若 $y' \in \Psi(x, y, d)$, 并且不存在 $y \in \Psi(x, y, d)$, 使得

$$y' - y \in \text{int}K,$$

则称 y' 是集值映射 ϕ 在点 (x^0, y^0) 处沿方向 d 的 CK -弱方向导数. ϕ 在点 (x^0, y^0) 处沿 d 的所有 CK -弱方向导数组成的集合记作

$\phi_w^c(x^0, y^0; d)$, 即 $\phi_w^c(x^0, y^0; d) = \mathcal{E}_w(\Psi(x^0, y^0, d), K)$.

注 7.2.1 设 \mathcal{B} 是 Banach 空间, $f: \mathcal{B} \rightarrow R \cup \{+\infty\}$ 是广义实值函数, 并且 $K = R_+$, $x^0 \in \mathcal{B}$, $y^0 = f(x^0)$, $d \in \mathcal{B}$. 由定义 7.2.1 我们有

$$f_w^c(x^0, y^0; d) = f^c(x^0; d) = \inf\{\eta \mid \eta \in \Psi(x^0, y^0, d)\},$$

其中 $f^c(x^0; d)$ 是 f 的 C -切导数. 根据定理 5.4.2, 当 f 在 x^0 处是 Lipschitz 的, 则有

$$f_w^c(x^0, y^0; d) = \limsup_{\substack{x \rightarrow x^0 \\ t \rightarrow 0^+}} \frac{f(x + td) - f(x)}{t} = f^0(x^0; d).$$

可见, CK-弱方向导数是一般函数的 G -方向导数的推广. 记

$$\hat{D} = \{d \in \mathcal{X} \mid \Psi(x^0, y^0, d) \neq \emptyset\},$$

$$K\text{-epi } \Psi = \{(x, y) \mid x \in \mathcal{X}, y \in \Psi(x^0, y^0, x) + K\}.$$

定义 7.2.2 设点 $x^0 \in \mathcal{X}$, $y^0 \in \phi(x^0)$, $p \in \text{int}K$, 则称集合 $\partial_w^c \phi(x^0, y^0)_p$,

$$= \left\{x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{x \in \hat{D}} [\phi_w^c(x^0, y^0; d) - \langle x^*, d \rangle p], K \right) \right\}$$

是集值映射 ϕ 在点 (x^0, y^0) 处关于 p 的 CK-弱次微分, 称 $x^* \in \partial_w^c \phi(x^0, y^0)_p$ 是 ϕ 在点 (x^0, y^0) 处关于 p 的 CK-弱次梯度. 若 $\partial_w^c \phi(x^0, y^0)_p \neq \emptyset$, 则称 ϕ 在点 (x^0, y^0) 处关于 p 是 CK-弱次可微的.

定理 7.2.1 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$. 若 $\Psi(x, y, \theta)$ 是 K -有界的, 则

$$\phi_w^c(x, y; d) \neq \emptyset \quad \forall d \in \mathcal{X},$$

$$\partial_w^c \phi(x, y)_p$$

$$= \left\{x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{X}} [\Psi(x, y, d) - \langle x^*, d \rangle p], K \right) \right\}$$

$$= \left\{ x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{(d, r) \in T_{K\text{-epi}\psi}^c(x, y)} [r - \langle x^*, d \rangle p], K \right) \right\}.$$

证明 取任意的 $d^1, d^2 \in \mathcal{D}$, 任意的 $y^1 \in \Psi(x, y, d^1)$ 和任意的 $y^2 \in \Psi(x, y, d^2)$, 由定义 7.2.1 可知 $(d^1, y^1), (d^2, y^2) \in T_{K\text{-epi}\psi}^c(x, y)$. 因为 $T_{K\text{-epi}\psi}^c(x, y)$ 是凸锥, $(d^1 + d^2, y^1 + y^2) \in T_{K\text{-epi}\psi}^c(x, y)$, 由 (7.2.1) 我们有 $y^1 + y^2 \in \Psi(x, y, d^1 + d^2)$, 因此得

$$\Psi(x, y, d^1) + \Psi(x, y, d^2) \subset \Psi(x, y, d^1 + d^2). \quad (7.2.2)$$

取任意的 $d \in \mathcal{D}$, 令 $b \in \Psi(x, y, -d)$, 根据 (7.2.2) 有

$$\Psi(x, y, d) + b \subset \Psi(x, y, \theta).$$

因为 $\Psi(x, y, \theta)$ 是 K -有界的, 所以存在 $a \in \mathcal{D}$ 使得 $\Psi(x, y, \theta) \subset a + K$, 因此

$$\Psi(x, y, d) \subset a - b + K.$$

根据定义 6.5.2 的 (1) 知 $\Psi(x, y, d)$ 也是 K -有界的. 因为 $T_{K\text{-epi}\psi}^c(x, y)$ 是闭的, K 是闭的, 所以 $\Psi(x, y, d)$ 是 K -闭集. 据此, 由引理 6.5.1 知 $\Psi(x, y, d)$ 是 K -半紧的. 于是, 根据引理 6.2.5 得

$$\phi_w^c(x, y; d) \neq \emptyset \quad \forall d \in \mathcal{D}.$$

再由引理 6.5.2 便知

$$\Psi(x, y, d) \subset \phi_w^c(x, y, d) + K. \quad (7.2.3)$$

设集合

$$S_1 = \left\{ x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{x \in \mathcal{X}} [\Psi(x, y, d) - \langle x^*, d \rangle p], K \right) \right\},$$

$$S_2 = \left\{ x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{(d, r) \in T_{K\text{-epi}\psi}^c(x, y)} [r - \langle x^*, d \rangle p], K \right) \right\}.$$

显然有 $S_1 = S_2$, 因此仅需证明 $S_1 = \partial_{\omega}^c \phi(x, y)_p$. 若 $x^* \in S_1$, 则不存在 $y' \in \Psi(x, y, d)$ 使得

$$0 - (y' - \langle x^*, d \rangle p) \in \text{int}K.$$

由此, 按定义 7.2.1 知 $\phi_{\omega}^c(x, y; d) \subset \Psi(x, y, d)$, 于是不存在 $y'' \in \phi_{\omega}^c(x, y; d)$ 使

$$0 - (y'' - \langle x^*, d \rangle p) \in \text{int}K.$$

因此, 由定义 7.2.2 得知 $x^* \in \partial_{\omega}^c \phi(x, y)_p$.

现设 $x^* \in \partial_{\omega}^c \phi(x, y)_p$, 反之假若 $x^* \notin S_1$, 则存在 $\bar{y} \in \Psi(x, y, d)$, 使得

$$0 - (\bar{y} - \langle x^*, d \rangle p) \in \text{int}K.$$

据此, 由 (7.2.3), 存在 $q' \in K$ 和 $\bar{y}' \in \phi_{\omega}^c(x, y; d)$, 使得 $\bar{y} = \bar{y}' + q'$, 从而

$$0 - (\bar{y}' + q' - \langle x^*, d \rangle p) \in \text{int}K,$$

即 $0 - (\bar{y}' - \langle x^*, d \rangle p) \in \text{int}K$. 于是由定义 7.2.2 可知 $x^* \notin \partial_{\omega}^c \phi(x, y)_p$, 这导致与已设 $x^* \in \partial_{\omega}^c \phi(x, y)_p$ 矛盾. \square

下面给出锥弱次微分的存在性定理.

定理 7.2.2 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$, $\Psi(x, y, \theta) \subset K$, $\text{int}(K\text{-epi}\Psi) \neq \emptyset$ (或 $\Psi(x, y, \cdot)$ 在某个 $\bar{d} \in \mathcal{X}$ 处是 $(-K)$ -有界的). 若存在 $d^0 \in \mathcal{X}$ 和 $q_0^* \in K^* \setminus \{0\}$ 使得 $\inf \langle q_0^*, \Psi(x, y, d^0) \rangle > -\infty$, 则存在线性泛函 $x^* \in \mathcal{X}^*$ 使得

$$\inf \langle q_0^*, \Psi(x, y, d^0) \rangle = \langle x^*, d^0 \rangle \langle q_0^*, p \rangle,$$

并且 $x^* \in \partial_{\omega}^c \phi(x, y)_p$.

证明 因为 $T_{K\text{-epi}\Psi}^c(x, y)$ 是凸锥, 所以对任意的 $\lambda > 0$ 和 $(d, r) \in T_{K\text{-epi}\Psi}^c(x, y)$ 有 $(\lambda d, \lambda r) \in T_{K\text{-epi}\Psi}^c(x, y)$. 由 (7.2.1) 知 $\lambda r \in \Psi(x, y, \lambda d)$, 故 $\Psi(x, y, \lambda d) = \lambda \Psi(x, y, d)$. 显然, 我们有

$$\Psi(x, y, d^1) + \Psi(x, y, d^2) \subset \Psi(x, y, d^1 + d^2)$$

$$\forall d^1, d^2 \in \mathcal{X}.$$

因此, $\Psi(x, y, d)$ 关于 d 是 K -次线性的. 又显然有 $0 \in \Psi(x, y, 0)$, 于是由定理 7.1.5 和定理 7.2.1 得到结论. \square

定理 7.2.3 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$. 若 $\Psi(x, y, 0)$ 是 K -有界的, 则 $x^* \in \partial_{\omega}^c \phi(x, y)$, 当且仅当存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall d \in \mathcal{X}, \forall y' \in \phi_{\omega}^c(x, y; d). \quad (7.2.4)$$

证明 必要性. 设 $x^* \in \partial_{\omega}^c \phi(x, y)$, 作集合

$$E = \{y \in \mathcal{Y}^* \mid y = -r + \langle x^*, d \rangle p - q,\$$

$$(d, r) \in T_{K\text{-epi}\phi}^c(x, y), q \in K\}.$$

易知 E 是凸集, 现在证明 $\text{int}K \cap E = \emptyset$. 反之, 假若有 $q' \in \text{int}K \cap E$, 即 $-r + \langle x^*, d \rangle p - q = q'$, 则有 $-r + \langle x^*, d \rangle p = q + q' \in \text{int}K$. 由定理 7.2.1 得 $x^* \notin \partial_{\omega}^c \phi(x, y)$, 导致与已设 $x^* \in \partial_{\omega}^c \phi(x, y)$ 矛盾. 据此, 由定理 1.3.7 可知, 存在 $q^* \in \mathcal{Y}^* \setminus \{0\}$ 使得

$$\langle q^*, q \rangle \geq \langle q^*, -r + \langle x^*, d \rangle p - q' \rangle$$

$$\forall q \in K, \forall (d, r) \in T_{K\text{-epi}\phi}^c(x, y), q' \in K. \quad (7.2.5)$$

在 (7.2.5) 中令 $r = 0, d = 0, q' = 0$ 得 $\langle q^*, q \rangle \geq 0 (\forall q \in K)$, 故有 $q^* \in K^* \setminus \{0\}$. 在 (7.2.5) 中再令 $q' = 0, q = 0$, 得到

$$\langle q^*, r \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall (d, r) \in T_{K\text{-epi}\phi}^c(x, y).$$

由定义 7.1.1 和 $\phi_{\omega}^c(x, y; d) \subset \Psi(x, y, d)$ 即得 (7.2.4).

充分性. 现设 (7.2.4) 成立, 反之, 假设 $x^* \notin \partial_{\omega}^c \phi(x, y)$. 由定义 7.2.2 知存在 $y' \in \phi_{\omega}^c(x, y; d)$, 使得

$$0 - (y' - \langle x^*, d \rangle p) \in \text{int}K.$$

由 $q^* \in K^* \setminus \{0\}$ 有

$$\langle q^*, -y' + \langle x^*, d \rangle p \rangle > 0,$$

这导致与(7.2.4)矛盾. \square

推论 7.2.4 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$, $p^1, p^2 \in \text{int}K$. 若 $\Psi(x, y, \theta)$ 是 K -有界的, 并且 $x^* \in \partial_\omega^C \phi(x, y)_{p^1}$, 则存在 $q^* \in K^* \setminus \{0\}$, 使得 $\alpha x^* \in \partial_\omega^C \phi(x, y)_{p^2}$, 其中 $\alpha = \langle q^*, p^1 \rangle / \langle q^*, p^2 \rangle$.

证明 由定理 7.2.3 容易推得证明. \square

推论 7.2.4 说明了 ϕ 的 K -次梯度与 p 无关.

定理 7.2.5 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$. 若 $\Psi(x, y, \theta)$ 是 K -有界的, 并且 $\partial_\omega^C \phi(x, y)_p \neq \emptyset$, 则 $\partial_\omega^C \phi(x, y)_p$ 是闭凸集.

证明 先证闭性. 用反证法, 假设有序列 $\{x_k^*\} \subset \partial_\omega^C \phi(x, y)_p$, $x_k^* \rightarrow x^*$, 但 $x^* \notin \partial_\omega^C \phi(x, y)_p$, 则由定理 7.2.1 知, 存在 $(d, r) \in T_{K-\text{epi}\phi}^C(x, y)$ 使得

$$- (r - \langle x^*, d \rangle p) \in \text{int}K.$$

因为 $-r + \langle x_k^*, d \rangle p \rightarrow -r + \langle x^*, d \rangle p (k \rightarrow \infty)$, 所以存在充分大的 k' , 使得

$$- (r - \langle x_{k'}^*, d \rangle p) \in \text{int}K.$$

再由定理 7.2.1 得知 $x_{k'}^* \notin \partial_\omega^C \phi(x, y)_p$, 这导致与假设矛盾.

现证凸性. 设 $x_1^*, x_2^* \in \partial_\omega^C \phi(x, y)_p$, 由定理 7.2.3 可知, 存在 $q_1^*, q_2^* \in K^* \setminus \{0\}$ 使得

$$\langle q_1^*, r \rangle \geq \langle x_1^*, d \rangle \langle q_1^*, p \rangle \quad \forall (d, r) \in T_{K-\text{epi}\phi}^C(x, y), \quad (7.2.6)$$

$$\langle q_2^*, r \rangle \geq \langle x_2^*, d \rangle \langle q_2^*, p \rangle \quad \forall (d, r) \in T_{K-\text{epi}\phi}^C(x, y). \quad (7.2.7)$$

对任意的 $\lambda \in (0, 1)$, 设 $x_\lambda^* = \lambda x_1^* + (1 - \lambda)x_2^*$, 我们要证明 $x_\lambda^* \in \partial_\omega^c \phi(x, y)_p$. 反之, 假设 $x_\lambda^* \notin \partial_\omega^c \phi(x, y)_p$, 则存在 $(d, r) \in T_{K-\text{epi}\phi}^c(x, y)$ 使得 $-r + \langle x_\lambda^*, d \rangle p \in \text{int}K$. 由 $q_1^*, q_2^* \in K^* \setminus \{0\}$ 和定理 1.4.11, 可推得

$$\langle q_1^*, -r + \langle x_\lambda^*, d \rangle p \rangle > 0, \quad (7.2.8)$$

$$\langle q_2^*, -r + \langle x_\lambda^*, d \rangle p \rangle > 0. \quad (7.2.9)$$

考虑 (7.2.6) + (7.2.8) 和 (7.2.7) + (7.2.9), 可得

$$\langle x_\lambda^*, d \rangle > \langle x_1^*, d \rangle, \quad (7.2.10)$$

$$\langle x_\lambda^*, d \rangle > \langle x_2^*, d \rangle. \quad (7.2.11)$$

从 $\lambda(7.2.10) + (1 - \lambda) \times (7.2.11)$, 则得

$$\langle x_\lambda^*, d \rangle > \langle x_\lambda^*, d \rangle,$$

导致矛盾. \square

定理 7.2.6 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$, $\Psi(x, y, 0) \subset K$. 若 $\text{int}(K - \text{epi}\Psi) \neq \emptyset$, $\partial_\omega^c(x, y)_p \neq \emptyset$, 则对任给的 $d \in \mathcal{X}$ 存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\begin{aligned} & \inf \langle q^*, \phi_\omega^c(x, y; d) \rangle \\ &= \max \{ \langle x^*, d \rangle \langle q^*, p \rangle \mid x^* \in \partial_\omega^c \phi(x, y)_p \}. \end{aligned}$$

证明 任给 $d^1 \in \mathcal{X}$, 设 $\bar{x}^* \in \partial_\omega^c \phi(x, y)_p$, 使得

$$\langle \bar{x}^*, d^1 \rangle \geq \langle x^*, d^1 \rangle \quad \forall x^* \in \partial_\omega^c \phi(x, y)_p. \quad (7.2.12)$$

由定理 7.2.3, 对于 $\bar{x}^* \in \partial_\omega^c \phi(x, y)_p$ 存在 $q^* \in K^* \setminus \{0\}$ 使

$$\langle q^*, y' \rangle \geq \langle \bar{x}^*, d \rangle \langle q^*, p \rangle \quad \forall d \in \mathcal{X}, \forall y' \in \phi_\omega^c(x, y; d).$$

从上式显然有

$$\inf \langle q^*, \phi_\omega^c(x, y; d^1) \rangle \geq \langle \bar{x}^*, d^1 \rangle \langle q^*, p \rangle.$$

于是,由定理 7.2.2,存在 $x_1^* \in \partial_w^c \psi(x, y)$, 使得

$$\inf \langle q^*, \psi_w^c(x, y; d^1) \rangle = \langle x_1^*, d^1 \rangle \langle q^*, p \rangle,$$

因此有 $\langle x_1^*, d^1 \rangle \geq \langle \bar{x}^*, d^1 \rangle$. 再由(7.2.12)得

$$\langle x_1^*, d^1 \rangle = \langle \bar{x}^*, d^1 \rangle,$$

即定理结论成立. \square

推论 7.2.7 设 \mathcal{B} 是 Banach 空间. 若实值函数 $f: \mathcal{B} \rightarrow R$ 在 $x \in \mathcal{B}$ 处是 Lipschitz 的, $y = f(x)$, $p = 1$, $K = R_+$, 则

$$f^0(x, y; d) = \max \{ \langle x^*, d \rangle \mid x^* \in \partial^0 f(x) \} \quad \forall d \in \mathcal{B}.$$

证明 由注 7.2.1 和定理 7.2.6, 即知结论成立. \square

现在,讨论锥凸集值映射的锥弱次微分与第 6 章的锥次微分之间的关系.

定理 7.2.8 设 $x \in \mathcal{X}$, $y \in \psi(x)$, 集值映射 ψ 在 \mathcal{X} 上是 K -凸的, $\hat{D} = \mathcal{X}$. 若 $\Psi(x, y, \theta)$ 是 K -有界的, 则

$$\begin{aligned} \partial_w^c \psi(x, y) \supset \left\{ x^* \in \mathcal{X}^* \mid y - \langle x^*, x \rangle p \right. \\ \left. \in \mathcal{E}_w \left(\bigcup_{x' \in \mathcal{X}} [\psi(x') - \langle x^*, x' \rangle p], K \right) \right\}. \end{aligned}$$

当 \mathcal{X} 和 \mathcal{Y} 是 Banach 空间时, 上式成为等式.

证明 由 ψ 是 K -凸的, 集合

$$C_{K\text{-epi}\psi}(x, y) = \text{cl} \left(\bigcup_{\lambda > 0} \lambda [K\text{-epi}\psi - (x, y)] \right)$$

非空, 由定义 5.2.2 知

$$T_{K\text{-epi}\psi}^c(x, y) \subset C_{K\text{-epi}\psi}(x, y).$$

若 \mathcal{X} 和 \mathcal{Y} 是 Banach 空间, 则由定理 5.2.2 知上式为等式. 由此, 根据定理 6.5.3 和定义 7.2.2 可得知本定理的结论成立. \square

定义 7.2.3 设点 $x^0 \in \mathcal{X}$, $y^0 \in \psi(x^0)$. 若存在 x^0 的邻域 $U(x^0)$ 使得

$$y^0 \in \mathcal{E}_w(\phi(U(x^0)), K),$$

即不存在 $x \in U(x^0)$ 和 $y \in \phi(x)$ 使得

$$y^0 - y \in \text{int}K,$$

则称点 (x^0, y^0) 是集值映射 ϕ 的局部弱有效解.

定理 7.2.9 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\hat{D} = \mathcal{X}$. 若 $\Psi(x, y, \theta)$ 是 K -有界的, (x, y) 是 ϕ 的局部弱有效解, 则 $\theta \in \partial_w^c \phi(x, y)_r$.

证明 用反证法: 假设 $\theta \notin \partial_w^c \phi(x, y)_r$. 由定理 7.2.1 可知, 存在 $r \in \Psi(x, y, d)$ 使得

$$-r + \langle \theta, d \rangle p \in \text{int}K. \quad (7.2.13)$$

由此, 对任何收敛于 $(x, y, 0)$ 的序列 $\{(x^k, y, t_k)\}$ 存在收敛于 (d, r) 的序列 $\{(d^k, r^k)\}$, 使得

$$y + t_k r^k \in \phi(x^k + t_k d^k) + K, \quad k = 1, 2, \dots.$$

于是, 存在 $y^k \in \phi(x^k + t_k d^k)$ 和 $q^k \in K$, 使得

$$y + t_k r^k = y^k + q^k, \quad k = 1, 2, \dots,$$

从而有 $r^k = \frac{y^k - q^k - y}{t_k} \rightarrow r$. 又因为 $x^k + t_k r^k \rightarrow x$, 故存在充分大的 k' , 使 $x^{k'} + t_{k'} r^{k'} \in U(x)$, $y^{k'} \in \phi(x^{k'} + t_{k'} r^{k'})$, 以及 $-\frac{y^{k'} - q^{k'} - y}{t^{k'}} \in \text{int}K$, 或即

$$y - y^{k'} \in \text{int}K.$$

据此, 由定义 7.2.3 得到 $y \notin \mathcal{E}_w(U(x), K)$, 这与已知矛盾. 定理得证. \square

§ 7.3 G-锥弱次微分

本节先介绍一种 G -切锥, 它比 C -切锥小. 从这种 G -切锥出

发,我们引进集值映射的 G -锥弱方向导数和 G -锥弱次微分.

设 \mathcal{X} 和 \mathcal{Y} 是局部凸线性拓扑空间, $\psi: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \psi(x)$ 是集值映射, $K \subset \mathcal{Y}$ 是内部非空的尖闭凸锥.

定义 7.3.1 设点 $x^0 \in \mathcal{X}$, $y^0 \in \mathcal{Y}$. 我们称集合 $G_{K\text{-epi}\psi}(x^0, y^0)$ 是实值映射 ψ 在 \mathcal{X} 上的 K -上图象 $K\text{-epi}\psi$ 在点 (x^0, y^0) 处的 G -切锥, 即对任意的 $(d, r) \in G_{K\text{-epi}\psi}(x^0, y^0) \subset \mathcal{X} \times \mathcal{Y}$, 若对任何收敛于 (x^0, y^0) 的序列 $\{(x^k, y^k)\} \subset K\text{-epi}\psi$ 和收敛于 0 的正数列 $\{t_k\} (k = 1, 2, \dots)$, 存在一个收敛于 r 的序列 $\{r^k\} \subset \mathcal{Y}$, 使得

$$(x^k + t_k d, y^k + t_k r^k) \in K\text{-epi}\psi, k = 1, 2, \dots.$$

关于 G -切锥有如下性质.

定理 7.3.1 设 $x \in \mathcal{X}$, $y \in \psi(x)$.

(1) $(0, 0) \in G_{K\text{-epi}\psi}(x, y) \subset T_{K\text{-epi}\psi}^{\mathcal{C}}(x, y)$.

(2) $G_{K\text{-epi}\psi}(x, y)$ 是凸锥.

(3) 若对任何 $(d, r^k) \in G_{K\text{-epi}\psi}(x, y) (k = 1, 2, \dots)$ 有 $(d, r^k) \rightarrow (d, r)$, 则 $(d, r) \in G_{K\text{-epi}\psi}(x, y)$.

证明 (1) 由定义 7.3.1 容易直接推得.

(2) 设 $(d, r) \in G_{K\text{-epi}\psi}(x, y)$, $\lambda > 0$, 按定义 7.3.1, 则对任何收敛于 (x, y) 的序列 $\{(x^k, y^k)\} \subset K\text{-epi}\psi$ 和收敛于 0 的正数列 $\{t_k\}$, 有收敛于 r 的序列 $\{r^k\} \subset \mathcal{Y}$, 使得

$$(x^k + \lambda t_k d, y^k + \lambda t_k r^k) \in K\text{-epi}\psi, k = 1, 2, \dots.$$

由于 $\lambda r^k \rightarrow \lambda r$, 据定义 7.3.1 得

$$(\lambda d, \lambda r) \in G_{K\text{-epi}\psi}(x, y),$$

因此 $G_{K\text{-epi}\psi}(x, y)$ 是一个锥.

为证 $G_{K\text{-epi}\psi}(x, y)$ 是凸的, 设 $(d^1, r^1), (d^2, r^2) \in G_{K\text{-epi}\psi}(x, y)$. 对任何收敛于 $(x^k, y^k, 0)$ 的序列 $\{(x^k, y^k, t_k)\} \subset K\text{-epi}\psi \times (0, +\infty)$, 则存在收敛于 r^1 的序列 $\{r_k^1\}$, 使得

$$(x^k + t_k d^1, y^k + t_k r_k^1) \in K\text{-epi}\psi, k = 1, 2, \dots.$$

因为 $x^k + t_k d^1 \rightarrow x$, $y^k + t_k r_k^1 \rightarrow y$, $t_k \rightarrow 0 (k \rightarrow \infty)$, $(d^2, r^2) \in G_{K\text{-epi}\psi}(x, y)$, 所以存在收敛于 r^2 的序列 $\{r_k^2\}$, 使得

$$(x^k + t_k d^1 + t_k d^2, y^k + t_k r_k^1 + t_k r_k^2) \in K\text{-epi}\psi, \quad k = 1, 2, \dots.$$

由于 $r_k^1 + r_k^2 \rightarrow r^1 + r^2$, 故由上式可推得

$$(d^1 + d^2, r^1 + r^2) \in G_{K\text{-epi}\psi}(x, y),$$

于是证明了 $G_{K\text{-epi}\psi}(x, y)$ 是凸的.

(3) 考虑任何收敛于 $(x, y, 0)$ 的序列 $\{(x^k, y^k, t_n)\} \subset K\text{-epi}\psi \times (0, +\infty) (n = 1, 2, \dots)$. 由 $(d, r^k) \in G_{K\text{-epi}\psi}(x, y) (k = 1, 2, \dots)$, 则对于每一个 $k (= 1, 2, \dots)$ 存在序列 $\{r_n^k\}$ 收敛于 $r^k (n \rightarrow \infty)$, 使得

$$(x^k + t_n d, y^k + t_n r_n^k) \in K\text{-epi}\psi, \quad n = 1, 2, \dots; \quad k = 1, 2, \dots.$$

设 V 是 \mathscr{Y} 的原点的任意邻域, 则存在另一个原点邻域 $U \subset \mathscr{Y}$, 使得 $U + U \subset V$. 因为 $r^k \rightarrow r$, 所以存在 $N > 0$, 使得 $r^k \in r + U$ 对任意的 $k \geq N$ 成立. 又因为 $r_n^k \rightarrow r^k (n \rightarrow \infty)$ 对于每一个 $k (= 1, 2, \dots)$ 成立, 故对于每一个 $k (= 1, 2, \dots)$ 均存在 $N_k > 0$, 使得

$$r_n^k \in r^k + U, \quad n \geq N_k,$$

从而得

$$r_n^k \in r + U + U \subset r + V, \quad n \geq N_k, \quad k \geq N.$$

对每一个 $k (\geq N)$ 对应地取一个 $r_n = r_n^k (n = 1, 2, \dots)$, 故获得一个序列 $\{r_n\}$ 收敛于 r , 并且有

$$(x^n + t_n d, y^n + t_n r_n) \in K\text{-epi}\psi, \quad n = 1, 2, \dots,$$

即 $(d, r) \in G_{K\text{-epi}\psi}(x, y)$. \square

定义 7.3.2 设点 $x^0 \in \mathscr{X}$, $y^0 \in \psi(x^0)$, $d \in \mathscr{X}$. 令集合

$$\Psi^0(x^0, y^0, d) = \{r \in \mathscr{Y} \mid (d, r) \in G_{K\text{-epi}\psi}(x^0, y^0)\},$$

$$\bar{D} = \{d \in \mathscr{X} \mid \Psi^0(x^0, y^0, d) \neq \emptyset\}.$$

若 $y'(x^0, y^0; d) \in \Psi^0(x^0, y^0, d)$, 并且不存在 $y \in \Psi^0(x^0, y^0, d)$ 使得

$$y'(x^0, y^0, d) - y \in \text{int}K,$$

则称 $y'(x^0, y^0; d)$ 是集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 GK-弱方向导数. ψ 在点 (x^0, y^0) 处沿 d 的所有 GK-弱方向导数组成的集合记作 $\phi_w^0(x^0, y^0; d)$, 即

$$\phi_w^0(x^0, y^0; d) = \mathcal{E}_w(\Psi^0(x^0, y^0, d), K).$$

定义 7.3.3 设点 $x^0 \in \mathcal{X}$, $y^0 \in \psi(x^0)$, $\bar{D} \neq \emptyset$, $p \in \text{int}K$. 称集合

$$\begin{aligned} & \partial_w^0 \psi(x^0, y^0)_p \\ &= \left\{ x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{d \in \bar{D}} [\phi_w^0(x^0, y^0; d) - \langle x^*, d \rangle p], K \right) \right\} \end{aligned}$$

是集值映射 ψ 在点 (x^0, y^0) 处关于 p 的 GK-弱次微分, 称 $x^* \in \partial_w^0 \psi(x^0, y^0)_p$ 是 ψ 在点 (x^0, y^0) 处关于 p 的 GK-弱次梯度. 若 $\partial_w^0 \psi(x^0, y^0)_p \neq \emptyset$, 则称 ψ 在点 (x^0, y^0) 处关于 p 是 GK-弱次可微的.

定理 7.3.2 设 $x \in \mathcal{X}$, $y \in \psi(x)$, $d \in \mathcal{X}$, $p \in \text{int}K$, $\bar{D} = \mathcal{X}$. 若 $\Psi^0(x, y, \theta)$ 是 K -有界的, 则

$$\phi_w^0(x, y; d) \neq \emptyset \quad \forall d \in \mathcal{X},$$

$$\partial_w^0 \psi(x, y)_p$$

$$\begin{aligned} &= \left\{ x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{X}} [\phi_w^0(x, y; d) - \langle x^*, d \rangle p], K \right) \right\} \\ &= \left\{ x^* \in \mathcal{X}^* \mid \theta \in \mathcal{E}_w \left(\bigcup_{(d, r) \in G_{K\text{-epi}\psi}(x, y)} [r - \langle x^*, d \rangle p], K \right) \right\}. \end{aligned}$$

证明 与定理 7.2.1 的证明类似. \square

定理 7.3.3 设 $x \in \mathcal{X}$, $y \in \psi(x)$, $p \in \text{int}K$, $\bar{D} = \mathcal{X}$. 又设 $\Psi^0(x, y, \theta) \subset K$, $\text{int}(K\text{-epi}\Psi^0) \neq \emptyset$ (或 $\Psi^0(x, y, d)$ 在某个 d 附近是 K -有界的). 若存在 $d^0 \in \mathcal{X}$ 和 $q_0^* \in K^* \setminus \{\theta\}$, 使得

$\inf \langle q_0^*, \Psi^0(x, y; d^0) \rangle > -\infty$, 则存在线性泛函 $x^* \in \mathcal{X}^*$ 使得 $\inf \langle q^*, \Psi^0(x, y; d^0) \rangle = \langle x^*, d^0 \rangle \langle q_0^*, p \rangle$ 和 $x^* \in \partial_\omega^0 \phi(x, y)_p$.

证明 由定理 7.2.1、定理 7.1.5 和定理 7.3.2 可以推得本定理. \square

定理 7.3.4 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $p \in \text{int}K$, $\bar{D} = \mathcal{X}$. 若 $\Psi^0(x, y, \theta)$ 是 K -有界的, $\partial_\omega^0 \phi(x, y)_p \neq \emptyset$, 则 $\partial_\omega^0 \phi(x, y)_p$ 是闭凸集.

证明 与定理 7.2.6 的证明类似. \square

定理 7.3.5 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $p \in \text{int}K$, $\bar{D} = \mathcal{X}$. 若 $\Psi^0(x, y, \theta)$ 是 K -有界的, 则 $x^* \in \partial_\omega^0 \phi(x, y)_p$ 当且仅当存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' \rangle \geq \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall d \in \mathcal{X}, \forall y' \in \phi_\omega^0(x, y; d).$$

证明 与定理 7.2.3 的证明类似. \square

推论 7.3.6 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\bar{D} = \mathcal{X}$, 并且 $p^1, p^2 \in \text{int}K$. 若 $\Psi^0(x, y, \theta)$ 是 K -有界的, $x^* \in \partial_\omega^0 \phi(x, y)_{p^1}$, 则存在 $q^* \in K^* \setminus \{0\}$ 使得 $\alpha x^* \in \partial_\omega^0 \phi(x, y)_{p^2}$, 其中 $\alpha = \langle q^*, p^1 \rangle / \langle q^*, p^2 \rangle$.

证明 由定理 7.3.5 可得证. \square

定理 7.3.7 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\bar{D} = \mathcal{X}$, $p \in \text{int}K$. 若 $\Psi^0(x, y, \theta) \subset K$, $\text{int}(K\text{-epi}\Psi^0) \neq \emptyset$, $\partial_\omega^0 \phi(x, y)_p \neq \emptyset$, 则对任给的 $d \in \mathcal{X}$ 存在一 $q^* \in K^* \setminus \{0\}$, 使得

$$\begin{aligned} & \text{int} \langle q^*, \phi_\omega^0(x, y; d) \rangle \\ &= \max \{ \langle x^*, d \rangle \langle q^0, p \rangle \mid x^* \in \partial_\omega^0 \phi(x, y)_p \}. \end{aligned}$$

证明 与定理 7.2.6 的证明类似. \square

定义 7.3.4 设 $\phi_1(x)$ 和 $\phi_2(x) (\forall x \in \mathcal{X})$ 是 \mathcal{Y} 中的非空集合, $\phi_1: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto \phi_1(x)$ 和 $\phi_2: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto \phi_2(x)$ 是集值映射, 定义集值映射 $\Theta: \mathcal{X} \times \mathcal{Y} \rightarrow 2^{\mathcal{Y}} \times 2^{\mathcal{Y}}, (x, y) \mapsto \Theta(x, y)$,

$$\Theta(x, y) = \{(y^1, y^2) \in \mathcal{Y} \times \mathcal{Y} \mid y^1 \in \phi_1(x),$$

$$y^2 \in \psi_2(x), y = y^1 + y^2\}. \quad (7.3.1)$$

若点列 $\{(x^k, y^k)\}$ 收敛于点 (x^0, y^0) , 则对任何 $(y_1^0, y_2^0) \in \Theta(x^0, y^0)$, 存在一个序列 $\{(y_k^1, y_k^2)\} \subset \Theta(x^k, y^k)$ 收敛于点 $(y_1^0, y_2^0) \in \Theta(x^0, y^0)$, 则称 $\Theta(x, y)$ 在点 (x^0, y^0) 处是局部有界的.

引理 7.3.8 设 $\psi_1(x)$ 和 $\psi_2(x) (\forall x \in \mathcal{X})$ 是 \mathcal{Y} 中的非空集合, $\psi_1: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto \psi_1(x)$ 和 $\psi_2: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto \psi_2(x)$ 是集值映射. 又设 $x \in \mathcal{X}, y \in \psi_1(x) + \psi_2(x)$. 若由 (7.3.1) 定义的集值映射 Θ 在点 (x, y) 处局部有界, 则

$$\begin{aligned} \bigcup_{(y^1, y^2) \in \Theta(x, y)} [(\psi_1)_w^0(x, y^1; d) + (\psi_2)_w^0(x, y^2; d)] \\ \subset (\psi_1 + \psi_2)_w^0(x, y; d), \end{aligned} \quad (7.3.2)$$

其中 $d \in \{d \in \mathcal{X} | (\psi_1 + \psi_2)_w^0(x, y; d) \neq \emptyset\}$.

证明 考虑任何收敛于 $(x, y, 0)$ 的序列

$$\{(x^k, y^k, t_k)\} \subset K\text{-epi}(\psi_1 + \psi_2) \times (0, +\infty) \quad (k = 1, 2, \dots).$$

对任何 $(y^1, y^2) \in \Theta(x, y)$, 设 $r^1 \in \Psi_1^0(x, y^1, d), r^2 \in \Psi_2^0(x, y^2, d)$, 即 $(d, r^1) \in G_{K\text{-epi}\psi_1}(x, y^1), (d, r^1) \in G_{K\text{-epi}\psi_2}(x, y^2)$, 由定义 7.3.3 可知, 存在序列 $\{(y_k^1, y_k^2)\} \subset \Theta(x^k, y^k)$ 收敛于 $(y^1, y^2) (k \rightarrow \infty)$, 其中 $y^k = y_k^1 + y_k^2, y = y^1 + y^2$. 因此, 由 G -切锥的定义知存在序列 $\{r_k^1\}$ 和 $\{r_k^2\}$, $r_k^1 \rightarrow r^1, r_k^2 \rightarrow r^2$, 使得

$$y_k^1 + t_k r_k^1 \in \psi_1(x^k + t_k d) + K,$$

$$y_k^2 + t_k r_k^2 \in \psi_2(x^k + t_k d) + K,$$

从而得

$$\begin{aligned} y_k^1 + y_k^2 + t_k r_k^1 + t_k r_k^2 &= y^k + t_k(r_k^1 + r_k^2) \in \psi_1(x^k + t_k d) + K \\ &\quad + \psi_2(x^k + t_k d) + K \subset (\psi_1 + \psi_2)(x^k + t_k d) + K. \end{aligned}$$

因为 $r_k^1 + r_k^2 \rightarrow r^1 + r^2$, 所以 $(d, r^1 + r^2) \in G_{K\text{-epi}(\psi_1 + \psi_2)}(x, y)$,

故(7.3.2)成立. \square

定理 7.3.9 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\bar{D} = \mathcal{X}$, $p \in \text{int}K$. 若 $\Psi^0(x, y, \theta)$ 是 K -有界的, 则 $\phi_w^0(x, y; d)$ 关于 d 在 \mathcal{X} 上是 K -凸的.

证明 因为 $\Psi^0(x, y, \theta)$ 是 K -有界的, 可以推出对任何 $d \in \mathcal{X}$, $\Psi^0(x, y, d)$ 是 K -有界的. 由定理 7.3.1 可知 $\Psi^0(x, y, d)$ ($\forall d \in \mathcal{X}$) 是 K -闭的, 由引理 6.5.1 知 $\Psi^0(x, y, d)$ 是 K -半紧的, 再由引理 6.5.2 便得到

$$\Psi^0(x, y, d) \subset \phi_w^0(x, y; d) + K \quad \forall d \in \mathcal{X}. \quad (7.3.3)$$

又对任意的 $d^1, d^2 \in \mathcal{X}$ 和 $\lambda \in (0, 1)$, 由定理 7.3.1 有

$$\begin{aligned} & \lambda \Psi^0(x, y, d^1) + (1 - \lambda) \Psi^0(x, y, d^2) \\ & \in \Psi^0(x, y, \lambda d^1 + (1 - \lambda)d^2). \end{aligned} \quad (7.3.4)$$

从(7.3.3)和(7.3.4)则得

$$\begin{aligned} & \lambda \phi_w^0(x, y; d^1) + (1 - \lambda) \phi_w^0(x, y; d^2) \\ & \subset \lambda \Psi^0(x, y, d^1) + (1 - \lambda) \Psi^0(x, y, d^2) \\ & \subset \Psi^0(x, y, \lambda d^1 + (1 - \lambda)d^2) \\ & \subset \phi_w^0(x, y; \lambda d^1 + (1 - \lambda)d^2) + K. \quad \square \end{aligned}$$

定理 7.3.10 设 $\phi_1(x)$ 和 $\phi_2(x)$ ($\forall x \in \mathcal{X}$) 是 \mathcal{Y} 中的非空集合, $\phi_1: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi_1(x)$ 和 $\phi_2: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, $x \mapsto \phi_2(x)$ 是集值映射. 又设 $x \in \mathcal{X}$, $y \in \phi_1(x) + \phi_2(x)$, $\Psi_1^0(x, y, \theta)$ 和 $\Psi_2^0(x, y, \theta)$ 是 K -有界的, $\bar{D}_1 = \bar{D}_2 = \mathcal{X}$ ($\bar{D}_1 = \{d \mid (\phi_1)_w^0(x, y; d) \neq \emptyset\}$, $\bar{D}_2 = \{d \mid (\phi_2)_w^0(x, y; d) \neq \emptyset\}$), $\text{int}(K\text{-epi}\Psi_1^0) \neq \emptyset$ 或者 $\text{int}(K\text{-epi}\Psi_2^0) \neq \emptyset$. 若按(7.3.1)定义的 $\Theta(x, y)$ 在点 (x, y) 处是局部有界的, 则

$$\partial_w^0(\phi_1 + \phi_2)(x, y)_p$$

$$\subset \bigcup_{(y^1, y^2) \in \Theta(x, y)} (\partial_w^0 \phi_1(x, y^1)_p + \partial_w^0 \phi_2(x, y^2)_p).$$

证明 取 $x^* \in \partial_w^0(\phi_1 + \phi_2)(x, y)_p$, 不妨设 $\text{int}(K\text{-epi}\Psi_2^0) \neq \emptyset$. 对任意的 $(y^1, y^2) \in \Theta(x, y)$, 作集合

$$S_1 = \{(d, r) \in \mathcal{X} \times \mathcal{Y} \mid d \in \mathcal{X}, r^1 \in \Psi_1^0(x, y^1, d),$$

$$q^1 \in K, r = -r^1 - q^1 + \langle x^*, d \rangle p\},$$

$$S_2 = \{(d, r) \in \mathcal{X} \times \mathcal{Y} \mid d \in \mathcal{X}, r^2 \in \Psi_2^0(x, y^2, d),$$

$$q^2 \in K, r = r^2 + q^2\}.$$

由定理 7.3.1 可知 S_1 和 S_2 是非空凸集, 并且 $\text{int}S_2 \neq \emptyset$. 以下证明 $S_1 \cap \text{int}S_2 = \emptyset$. 事实上, 反之, 假若存在 $(\bar{d}, \bar{r}) \in S_1 \cap \text{int}S_2$, 我们有

$$\bar{r} = -r^1 - q^1 + \langle x^*, \bar{d} \rangle p,$$

并且存在 $\alpha > 0$, 使 $(\bar{d}, \bar{r} - \alpha p) \in S_2$. 因此, 有 $r^2 \in \Psi_2^0(x, y, \bar{d})$ 使得 $\bar{r} - \alpha p = r^2 + q^2$, 其中 $q^2 \in K$, 从而

$$-r^1 - q^1 + \langle x^*, \bar{d} \rangle p = r^2 + q^2 + \alpha p,$$

即

$$-(r^1 + r^2 - \langle x^*, \bar{d} \rangle p) = q^1 + q^2 + \alpha p \in \text{int}K.$$

据此, 由引理 7.3.8 得 $r^1 + r^2 \in (\Psi_1 + \Psi_2)^0(x, y, \bar{d})$. 从定理 7.3.2 知 $x^* \notin \partial_w^0(\phi_1 + \phi_2)(x, y)_p$, 这与已设矛盾. 现在, 利用定理 1.3.7 可知存在 $(\bar{x}^*, \bar{y}^*) \in \mathcal{X}^* \times \mathcal{Y}^*$ 和 $(\bar{x}^*, \bar{y}^*) \neq \theta$, 使

$$\langle \bar{x}^*, d^{S_1} \rangle + \langle \bar{y}^*, r^{S_1} \rangle \geq \beta \geq \langle \bar{x}^*, d^{S_2} \rangle + \langle \bar{y}^*, r^{S_2} \rangle$$

$$\forall (d^{S_1}, r^{S_1}) \in S_1, \forall (d^{S_2}, r^{S_2}) \in S_2. \quad (7.3.5)$$

因为 $(\theta, \theta) \in S_1 \cap S_2$, 所以从 (7.3.5) 得 $\beta = 0$. 假设 $\bar{y}^* = \theta$, 则 $\bar{x}^* \neq \theta$, 从 (7.3.5) 有

$$\langle \bar{x}^*, d \rangle \geq 0 \quad \forall d \in \mathcal{X},$$

从而得 $\bar{x}^* = 0$, 导致矛盾. 因此, 有 $\bar{y}^* \neq 0$. 在 (7.3.5) 中设 $d^{S_2} = 0$, $r^{S_2} = q \in K$, 得

$$\langle -\bar{y}^*, q \rangle \geq 0 \quad \forall q \in K,$$

从而 $-\bar{y}^* \in K^* \setminus \{0\}$. 于是, 由定理 1.4.11 得 $\langle -\bar{y}^*, p \rangle > 0$. 在 (7.3.5) 中令 $q^* = -\bar{y}^*$, $d^{S_2} = d$, $r^{S_2} = r^2$, $x_2^* = \bar{x}^* / \langle q^*, p \rangle$, 我们有

$$\langle q^*, r^2 \rangle \geq \langle x_2^*, d \rangle \langle q^*, p \rangle \quad \forall d \in \mathcal{X}, \forall r^2 \in \Psi_2^0(x, y_2, d).$$

由此, 根据定理 7.3.5 得到 $x_1^* \in \partial_\omega^0 \psi_2(x, y_1)_p$. 在 (7.3.5) 中再令 $q^* = -\bar{y}^*$, $d^{S_1} = d$, $r^{S_1} = -r^1 + \langle x^*, d \rangle p$, $x_1^* = x^* - x_2^*$, 则有

$$\langle q^*, r^1 \rangle \geq \langle x_1^*, d \rangle \langle q^*, p \rangle, \quad \forall d \in \mathcal{X},$$

$$\forall r^1 \in \Psi_1^0(x, y_1, d).$$

于是, 由定理 7.3.5 又得到 $x_1^* \in \partial_\omega^0 \psi_1(x, y_1)_p$. \square

定理 7.3.11 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $p \in \text{int}K$, $\bar{D} = \mathcal{X}$, 并且 $\Psi^0(x, y, 0)$ 是 K -有界的. 若 (x, y) 是 ϕ 的局部弱有效解, 则 $0 \in \partial_\omega^0 \phi(x, y)_p$.

证明 与定理 7.2.9 的证明类似. \square

§ 7.4 算子锥次微分

在前面, 我们都是通过线性泛函来引进集值映射的锥次微分的. 本节将从线性算子来定义集值映射的锥次微分.

设 \mathcal{X} 和 \mathcal{Y} 是局部凸线性拓扑空间, $L(\mathcal{X}, \mathcal{Y})$ 表示 \mathcal{X} 到 \mathcal{Y} 的所有连续线性算子组成的集合, $\phi: \mathcal{X} \rightarrow 2^\mathcal{Y}$, $x \mapsto \phi(x)$ 是集值映射, $K \subset \mathcal{Y}$ 是内部非空的尖闭凸锥.

定义 7.4.1 设集合 $S \subset \mathcal{X}$ 非空, 集值映射 ϕ 在 S 上是 K

凸的, 又点 $x^0 \in S, y^0 \in \psi(x^0)$.

(1) 若存在 $\mathcal{A} \in L(\mathcal{H}, \mathcal{Y})$ 使得

$$y^0 - \mathcal{A}(x^0) \in \mathcal{E}\left(\bigcup_{x \in S} [\psi(x) - \mathcal{A}(x)], K\right),$$

则称 \mathcal{A} 是集值映射 ψ 在点 (x^0, y^0) 处的算子 K -次梯度. ψ 在点 (x^0, y^0) 处的所有算子 K -次梯度组成的集合称为 ψ 在点 (x^0, y^0) 处的算子 K -次微分, 记作 $\partial\psi(x^0, y^0)_L$.

(2) 若存在 $\mathcal{A} \in L(\mathcal{H}, \mathcal{Y})$ 使得

$$y^0 - \mathcal{A}(x^0) \in \mathcal{E}_w\left(\bigcup_{x \in S} [\psi(x) - \mathcal{A}(x)], K\right),$$

则称 \mathcal{A} 是集值映射 ψ 在点 (x^0, y^0) 处的算子 K -弱次梯度. ψ 在点 (x^0, y^0) 处的所有算子 K -弱次梯度组成的集合称为 ψ 在点 (x^0, y^0) 处的算子 K -弱次微分, 记作 $\partial_w\psi(x^0, y^0)_L$.

由定义 6.1.3 和定义 7.4.1 易知有

$$\partial\psi(x^0, y^0)_L \subset \partial\psi(x^0, y^0)_L, \quad \partial_w\psi(x^0, y^0)_L \subset \partial_w\psi(x^0, y^0)_L. \quad (7.4.1)$$

下面是存在性定理.

定理 7.4.1 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\psi(x) \neq \emptyset (\forall x \in S)$, $x \in \text{int}S$, $p \in \text{int}K$. 又设 ψ 在 S 上是 K -凸的, $\text{int}(K\text{-epi}\psi) \neq \emptyset$, $y \in \mathcal{E}(\psi(x), K)$.

(1) $\partial_w\psi(x, y)_L \neq \emptyset$.

(2) 若 ψ 在 x 处是 K -严格凸的, 则 $\partial\psi(x, y)_L \neq \emptyset$.

证明 由定理 6.2.4 和 (7.4.1) 即可得证. \square

定理 7.4.2 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\psi(x) \neq \emptyset (\forall x \in S)$, $x \in S, y \in \psi(x)$. 若 ψ 在 S 是上 K -凸的, 则 $\mathcal{A} \in \partial_w\psi(x, y)_L$ 当且仅当存在 $q^* \in K^* \setminus \{0\}$, 使得

$$\langle q^*, y' - y \rangle \geq \langle q^*, \mathcal{A}(x') - \mathcal{A}(x) \rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

证明 与定理 6.2.11 的(2)和定理 6.2.12 的证明类似. \square

定理 7.4.3 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $\phi(x) \neq \emptyset$ ($\forall x \in S$), $x \in S, y \in \phi(x)$. 若 ϕ 在 S 上是 K -凸的, $\partial_w \phi(x, y)_L \neq \emptyset$, 则 $\partial_w \phi(x, y)_L$ 是闭凸集.

证明 与定理 6.3.1 和定理 6.3.2 的(1)的证明类似. \square

定理 7.4.4 设 $S \subset \mathcal{X}$ 是内部非空的凸集, $x \in S, \phi_1(x)$ 和 $\phi_2(x)$ ($\forall x \in S$) 非空,

$$\begin{aligned} (\phi_1 + \phi_2)(x) &= \{y^1 + y^2 \mid y^1 \in \phi_1(x), y^2 \in \phi_2(x), x \in S\}, \\ y^1 &\in \phi_1(x), y^2 \in \phi_2(x). \end{aligned}$$

$$\phi_1: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto \phi_1(x), \phi_2: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto \phi_2(x),$$

$\phi_1 + \phi_2: \mathcal{X} \rightarrow 2^{\mathcal{Y}}, x \mapsto (\phi_1 + \phi_2)(x)$ 是集值映射.

若 ϕ_1 和 ϕ_2 在 S 上是 K -凸的, ϕ_1 或 ϕ_2 在 x 处是连通的, 则

$$\partial_w(\phi_1 + \phi_2)(x, y^1 + y^2)_L \subset \partial_w \phi_1(x, y^1)_L + \partial_w \phi_2(x, y^2)_L.$$

证明 与定理 6.4.2 的证明类似. \square

定理 7.4.5 设 $S \subset \mathcal{X}$ 是非空凸集, $x \in S, y \in \phi(x)$, $\Psi(x, y, \theta)$ 和 $\psi'(x, y; d)$ 按定义 6.5.1 所确定. 若 ϕ 在 S 上是 K -凸的, $\hat{D} = \mathcal{X}$, $\Psi(x, y, \theta)$ 是 K -有界的, 则

$$\partial \phi(x, y)_L$$

$$\supset \left\{ \mathcal{A} \in L(\mathcal{X}, \mathcal{Y}) \mid \theta \in \mathcal{C} \left(\bigcup_{d \in \mathcal{X}} [\psi'(x, y; d) - \mathcal{A}(d)], K \right) \right\},$$

$$\partial_w \phi(x, y)_L$$

$$= \left\{ \mathcal{A} \in L(\mathcal{X}, \mathcal{Y}) \mid \theta \in \mathcal{C}_w \left(\bigcup_{d \in \mathcal{X}} [\psi'(x, y; d) - \mathcal{A}(d)], K \right) \right\}.$$

证明 与定理 6.5.3 的证明类似. \square

通过 C -切锥和 G -切锥, 也可以由算子构成集值映射的锥次微分.

定义 7.4.2 设点 $x^0 \in \mathcal{X}$, $y^0 \in \psi(x^0)$, $d \in \mathcal{X}$, $\phi_w^c(x^0, y^0; d)$ 是 ψ 在点 (x^0, y^0) 处沿方向 d 的 CK-弱方向导数, 则称集合

$$\partial_w^c \psi(x^0, y^0)_L$$

$$= \left\{ \mathcal{A} \in L(\mathcal{X}, \mathcal{Y}) \mid \mathbf{0} \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{X}} [\phi_w^c(x^0, y^0; d) - \mathcal{A}(d)], K \right) \right\}$$

是集值映射 ψ 在点 (x^0, y^0) 处的算子 CK-弱次微分, 称 $\mathcal{A} \in \partial_w^c \psi(x^0, y^0)_L$ 是 ψ 在点 (x^0, y^0) 处的算子 CK-弱次梯度. 若 $\partial_w^c \psi(x^0, y^0)_L \neq \emptyset$, 则称 ψ 在点 (x^0, y^0) 处是算子 CK-弱次可微的. 记

$$\partial_w^c \psi(x^0)_L = \{ \partial_w^c \psi(x^0, y^0)_L \mid y^0 \in \psi(x^0) \},$$

并称它是 ψ 在点 x^0 处的算子 CK-弱次微分.

定理 7.4.6 设 $x \in \mathcal{X}$, $y \in \psi(x)$, $\hat{D} = \mathcal{X}$. 若 $\Psi(x, y, \theta)$ 是 K -有界的, 则

$$\partial_w^c \psi(x, y)_L$$

$$= \left\{ \mathcal{A} \in L(\mathcal{X}, \mathcal{Y}) \mid \mathbf{0} \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{X}} [\Psi(x, y, d) - \mathcal{A}(d)], K \right) \right\}$$

$$= \left\{ \mathcal{A} \in L(\mathcal{X}, \mathcal{Y}) \mid \mathbf{0} \in \mathcal{E}_w \left(\bigcup_{(d, r) \in T_{K-\text{epi}\Psi}^c(x, y)} [r - \mathcal{A}(d)], K \right) \right\}.$$

证明 与定理 7.2.1 的证明类似. \square

定理 7.4.7 设 $x \in \mathcal{X}$, $y \in \psi(x)$, $p \in \text{int} K$, $\hat{D} = \mathcal{X}$. 又设 $\Psi(x, y, \theta) \subset K$, $\text{int}(K\text{-epi}\Psi) \neq \emptyset$. 若存在 $\bar{d} \in \mathcal{X}$ 和 $\bar{q}^* \in K^* \setminus \{\mathbf{0}\}$, 使得 $\inf \langle \bar{q}^*, \Psi(x, y, \bar{d}) \rangle > -\infty$, 则 $\partial_w^c \psi(x, y)_L \neq \emptyset$.

证明 由定理 7.2.2 知存在 $x^* \in \partial_w^c \psi(x, y)_p$. 作线性算子

$$\mathcal{A}: \mathcal{X} \rightarrow \mathcal{Y}, \mathcal{A}(x) = \langle x^*, x \rangle p \quad \forall x \in \mathcal{X},$$

则 $\mathcal{A} \in \partial_w^c \psi(x, y)_L$, 故得证. \square

定理 7.4.8 设 $x \in \mathcal{X}$, $y \in \psi(x)$, $\hat{D} = \mathcal{X}$. 又设 $\Psi(x, y,$

θ) 是 K -有界的.

(1) $\mathcal{A} \in \partial_w^c \phi(x, y)_L$ 当且仅当存在 $q^* \in K^* \setminus \{\theta\}$, 使得

$$\langle q^*, y' \rangle \geq \langle q^*, \mathcal{A}(d) \rangle \quad \forall d \in \mathcal{H}, \forall y' \in \phi_w^c(x, y; d).$$

(2) 若 $\partial_w^c \phi(x, y)_L \neq \emptyset$, 则 $\partial_w^c \phi(x, y)_L$ 是闭凸集.

(3) 若 (x, y) 是 ϕ 的局部弱有效解, 则 $\theta \in \partial_w^c \phi(x, y)_L$.

证明 (1) 与定理 7.2.3 的证明类似.

(2) 与定理 7.2.5 的证明类似.

(3) 与定理 7.2.9 的证明类似. \square

定义 7.4.3 设点 $x^0 \in \mathcal{H}$, $y^0 \in \phi(x^0)$, $d \in \mathcal{H}$. $\phi_w^0(x^0, y^0; d)$ 是集值映射 ϕ 在点 (x^0, y^0) 处沿 d 的 GK-弱方向导数, 则称集合

$$\begin{aligned} & \partial_w^0 \phi(x^0, y^0)_L \\ &= \left\{ \mathcal{A} \in L(\mathcal{H}, \mathcal{Y}) \mid \theta \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{H}} [\phi_w^0(x^0, y^0; d) - \mathcal{A}(d)], K \right) \right\} \end{aligned}$$

是集值映射 ϕ 在点 (x^0, y^0) 处的算子 GK-弱次微分, 称 $\mathcal{A} \in \partial_w^0 \phi(x^0, y^0)$ 是 ϕ 在点 (x^0, y^0) 处的算子 GK-弱次梯度. 若 $\partial_w^0 \phi(x^0, y^0)_L \neq \emptyset$, 则称 ϕ 在点 (x^0, y^0) 处是算子 GK-弱次可微的. 记 $\partial_w^0 \phi(x^0)_L = \{\partial_w^0 \phi(x^0, y^0)_L \mid y^0 \in \phi(x^0)\}$, 并称它是 ϕ 在点 x^0 处的算子 GK-弱次微分.

定理 7.4.9 设 $x \in \mathcal{H}$, $y \in \phi(x)$, $\bar{D} = \mathcal{H}$. 若 $\Psi^0(x, y, \theta)$ 是 K -有界的, 则

$$\begin{aligned} & \partial_w^0 \phi(x, y)_L \\ &= \left\{ \mathcal{A} \in L(\mathcal{H}, \mathcal{Y}) \mid \theta \in \mathcal{E}_w \left(\bigcup_{d \in \mathcal{H}} [\Psi^0(x, y, d) - \mathcal{A}(d)], K \right) \right\} \\ &= \left\{ \mathcal{A} \in L(\mathcal{H}, \mathcal{Y}) \mid \theta \in \mathcal{E}_w \left(\bigcup_{(d, r) \in G_{K, \text{epi}\phi}(x, y)} [r - \mathcal{A}(d)], K \right) \right\}. \end{aligned}$$

证明 与定理 7.3.2 的证明类似. \square

定理 7.4.10 设 $x \in \mathcal{H}$, $y \in \phi(x)$, $p \in \text{int}K$, $\bar{D} = \mathcal{H}$. 又

$\Psi^0(x, y, \theta) \subset K$, $\text{int}(K\text{-epi}\Psi^0) \neq \emptyset$. 若存在 $\bar{d} \in \mathcal{X}$ 和 $\bar{q}^* \in K^* \setminus \{\theta\}$, 使得 $\inf \langle \bar{q}^*, \phi_\omega^0(x, y; \bar{d}) \rangle > -\infty$, 则 $\partial_\omega^0 \phi(x, y)_L \neq \emptyset$.

证明 与定理 7.3.3 的证明类似. \square

定理 7.4.11 设 $x \in \mathcal{X}$, $y \in \phi(x)$, $\bar{D} = \mathcal{X}$. 又设 $\Psi^0(x, y, \theta)$ 是 K -有界的.

(1) 若 $\partial_\omega^0 \phi(x, y)_L \neq \emptyset$, 则 $\partial_\omega^0 \phi(x, y)_L$ 是闭凸集.

(2) $\mathcal{A} \in \partial_\omega^0 \phi(x, y)_L$ 当且仅当存在 $q^* \in K^* \setminus \{\theta\}$, 使得

$$\langle q^*, y' \rangle \geq \langle q^*, \mathcal{A}(d) \rangle \quad \forall d \in \mathcal{X}, \forall y' \in \phi_\omega^0(x, y; d).$$

(3) 若 (x, y) 是 ϕ 的局部弱有效解, 则 $\theta \in \partial_\omega^0 \phi(x, y)_L$.

证明 (1), (2) 和 (3) 依次与定理 7.3.4, 定理 7.3.5 和定理 7.3.11 的证明类似. \square

第 8 章 集值映射的切导数和切微分

这一章将进一步研究集值映射在其他意义下的导数和微分问题. 与前面介绍的次微分不同的是, 本章将借助相应图象的切锥或法锥, 来定义集值映射的导数和微分. 在本章的最后部分, 我们特别讨论有限维空间中集值映射的广义微分问题.

§ 8.1 切导数和切微分

设 \mathcal{B} 和 \mathcal{D} 是 Banach 空间, \mathcal{B}^* 和 \mathcal{D}^* 分别是 \mathcal{B} 和 \mathcal{D} 的对偶空间, $\psi: \mathcal{B} \rightarrow 2^{\mathcal{D}}$, $x \mapsto \psi(x)$ 是集值映射 (对于 $x \in \mathcal{B}$, 集合 $\psi(x) \neq \emptyset$, 且至少有一点 $x \in \mathcal{B}$ 对应的 $\psi(x)$ 是非单点集, 除非特别申明 ψ 是单值映射). 记 ψ 的有效域为

$$\text{dom } \psi = \{x \in \mathcal{B} \mid \psi(x) \neq \emptyset\},$$

ψ 的图象为

$$\text{graph } \psi = \{(x, y) \mid y \in \psi(x)\},$$

ψ 的象为

$$\text{Im } \psi = \bigcup_{x \in \mathcal{B}} \psi(x) = \bigcup_{x \in \text{dom } \psi} \psi(x).$$

又记 ψ 的逆集值映射为: $\psi^{-1}: \mathcal{D} \rightarrow 2^{\mathcal{B}}$, $y \mapsto \psi^{-1}(y)$. $x \in \psi^{-1}(y)$ 当且仅当 $y \in \psi(x)$.

显然有下面关系:

$$\text{dom } \psi^{-1} = \text{Im } \psi, \text{Im } \psi^{-1} = \text{dom } \psi.$$

这一节均设 $\text{graph } \psi$ 是 $\mathcal{B} \times \mathcal{D}$ 中的凸集, 点 $(x, y) \in \text{graph } \psi$, 即 $y = \psi(x)$. 显然, 集合 $\text{graph } \psi$ 在点 (x, y) 处的切锥 $T_{\text{graph } \psi}(x, y)$ 和法锥 $N_{\text{graph } \psi}(x, y)$ 是空间 $\mathcal{B} \times \mathcal{D}$ 中的闭凸集.

定义 8.1.1 设点 $x^0 \in \mathcal{B}$, $y^0 \in \psi(x^0)$, 考虑集值映射 $\phi^T(x^0, y^0): \mathcal{B} \rightarrow 2^{\mathcal{D}}$, $u \mapsto \phi^T(x^0, y^0)(u)$. 若对任意的 $v \in \phi^T(x^0, y^0)(u)$ 当且仅当

$$(u, v) \in T_{\text{graph } \psi}(x^0, y^0),$$

则称 $\phi^T(x^0, y^0)$ 是 ψ 在点 (x^0, y^0) 处的切导数. 考虑集值映射 $\partial^T \psi(x^0, y^0): \mathcal{D}^* \rightarrow 2^{\mathcal{B}^*}$, $q^* \mapsto \partial^T \psi(x^0, y^0)(q^*)$. 若对任意的 $p^* \in \partial^T \psi(x^0, y^0)(q^*)$ 当且仅当

$$(p^*, -q^*) \in N_{\text{graph } \psi}(x^0, y^0),$$

则称 $\partial^T \psi(x^0, y^0)$ 是 ψ 在点 (x^0, y^0) 处的切微分.

由定义 3.2.2 和定义 8.1.1 易得以下结果.

定理 8.1.1 设 $(x, y) \in \text{graph } \psi$, 则 $p^* \in \partial^T \psi(x, y)(q^*)$ 当且仅当

$$\langle q^*, y - y' \rangle \leq \langle p^*, x - x' \rangle \quad \forall x' \in \mathcal{B}, \quad \forall y' \in \psi(x'),$$

或

$$\langle p^*, u \rangle \leq \langle q^*, v \rangle \quad \forall u \in \mathcal{B}, \quad \forall v \in \phi^T(x, y)(u).$$

定理 8.1.2 设 $x \in \mathcal{B}$, $y \in \psi(x)$.

$$(1) (\phi^{-1})^T(y, x) = \phi^T(x, y)^{-1}.$$

$$(2) \partial^T(\phi^{-1})(y, x)(p^*) = -(\partial^T \psi(x, y))^{-1}(-p^*).$$

定理 8.1.3 设 $x^1, x^2 \in \mathcal{B}$, $y^1 \in \psi(x^1)$, $y^2 \in \psi(x^2)$, $p_1^* \in \partial^T \psi(x^1, y^1)(q_1^*)$, $p_2^* \in \partial^T \psi(x^2, y^2)(q_2^*)$, 则 $\langle q_1^* - q_2^*, y^1 - y^2 \rangle \leq \langle p_1^* - p_2^*, x^1 - x^2 \rangle$.

定理 8.1.4 设 $x \in \mathcal{B}$, $y \in \psi(x)$.

$$(1) \text{dom}(\phi^T(x, y)) \subset T_{\text{dom } \psi}(x).$$

$$(2) \text{Im}(\phi^T(x, y)) \subset T_{\text{Im } \psi}(y).$$

设 $K \subset \mathcal{B}$ 是内部非空的尖闭凸锥.

定理 8.1.5 设 $x \in \mathcal{B}$, $y \in \phi(x)$, $p \in \text{int}K$, $d \in \mathcal{B}$, 又 $\Psi(x, y, d)$ 由 (6.5.2) 确定, $\hat{D} = \mathcal{B}$, $\Psi(x, y, \theta)$ 是 K -有界的.

(1) $\phi^T(x, y)(d) \subset \Psi(x, y, d)$.

(2) 若 $x^* \in \partial^T \phi(x, y)(q^*)$, $q^* \in K^* \setminus \{\theta\}$, K^* 是 K 的对偶锥, 则

$$x^* (\langle q^*, p \rangle)^{-1} \in \partial_\omega \phi(x, y)_p.$$

(3) 若 $x^* \in \partial_\omega \phi(x, y)_p$, 则存在 $q^* \in K^* \setminus \{\theta\}$ 使得

$$\langle q^*, p \rangle x^* \in \partial^T \phi(x, y)(q^*).$$

证明 (1) 由定义 8.1.1 和定义 6.5.1 直接得证.

(2) 由定理 8.1.1 和定理 6.2.11 可以推得证明.

(3) 由定理 8.1.1 和定理 6.2.12 可以推证. \square

定理 8.1.6 设 $x \in \mathcal{B}$, $y \in \phi(x)$, 则 $v \in \phi^T(x, y)(u)$ 当且仅当

$$\liminf_{u' \rightarrow u, t > 0} d\left(v, \frac{\phi(x + tu') - y}{t}\right) = 0. \quad (8.1.1)$$

证明 因为 $(u, v) \in T_{\text{graph}\phi}(x, y)$ 当且仅当对任何 $\epsilon_1 > 0$, $\epsilon_2 > 0$, 存在 $u_{\epsilon_1} \in \mathcal{B}$, $v_{\epsilon_2} \in \mathcal{D}$, 满足 $\|u_{\epsilon_1}\| \leq \epsilon_1$, $\|v_{\epsilon_2}\| \leq \epsilon_2$ 和 $t_0 > 0$, 使得

$$(x + t(u + u_{\epsilon_1}), y + t(v + v_{\epsilon_2})) \in \text{graph } \phi \quad \forall t \in [0, t_0),$$

所以

$$v \in \frac{\phi(x + t(u + u_{\epsilon_1})) - y}{t} - v_{\epsilon_2}.$$

由此, 我们有

$$\inf_{0 < t < t_0} d\left(v, \frac{\phi(x + t(u + u_{\epsilon_1})) - y}{t}\right) \leq \epsilon_2. \quad (8.1.2)$$

由于 $\text{graph } \phi$ 是凸集, 对任何 $t_1 \leq t_2$, 有

$$\frac{\phi(x' + t_2 u') - y}{t_2} \subset \frac{\phi(x' + t_1 u') - y'}{t_1}$$

$$\forall x' \in \mathcal{D}, y' \in \phi(x'),$$

故

$$\lim_{t \rightarrow 0} d\left(v, \frac{\phi(x + tu') - y}{t}\right) = \inf_{t > 0} d\left(v, \frac{\phi(x + tu') - y}{t}\right).$$

由上式和(8.1.2)可得, 对 $\|u_{\epsilon_1}\| \leq \epsilon_1$ 有

$$\begin{aligned} & \inf_{0 < t < t_0} d\left(v, \frac{\phi(x + t(u + u_{\epsilon_1})) - y}{t}\right) \\ &= \inf_{t > 0} d\left(v, \frac{\phi(x + t(u + u_{\epsilon_1})) - y}{t}\right) \\ &= \lim_{t \rightarrow 0^+} d\left(v, \frac{\phi(x + t(u + u_{\epsilon_1})) - y}{t}\right). \end{aligned}$$

于是, 从上式得知

$$\inf_{\|u' - u\| \leq \epsilon_1} \inf_{t > 0} d\left(v, \frac{\phi(x + t(u + u_{\epsilon_1})) - y}{t}\right) \leq \epsilon_2.$$

令 $\epsilon_1 \rightarrow 0$, $\epsilon_2 \rightarrow 0$, 即得(8.1.1). \square

定理 8.1.7 设 $x^0, x \in \text{dom } \phi$, 则

$$\phi(x^0) - y \subset \phi^T(x, y)(x^0 - x), \quad \forall y \in \phi(x).$$

证明 由于 ϕ 在 $\text{dom } \phi$ 上是凸的, 故对任意的 $\lambda \in (0, 1)$ 和 $y^0 \in \phi(x^0)$, 有

$$(1 - \lambda)y + \lambda y^0 \subset \phi(x + \lambda(x^0 - x)),$$

即

$$y^0 - y \in \frac{\phi(x + \lambda(x^0 - x)) - y}{\lambda}.$$

由定义 8.1.1 得到 $y^0 - y \in \phi^T(x, y)(x^0 - x)$, 从而定理得证. \square

定义 8.1.2 设集合 $X \subset \mathcal{D}$ 非空, $\phi: X \rightarrow 2^{\mathcal{D}}$, $x \mapsto \phi(x)$ 是集

值映射, 点 $x^0 \in X$, $K \subset \mathscr{D}$ 是闭凸锥. 若对任意的 $y^0 \in \phi(x^0)$ 有 $\phi(x) \subset y^0 + K (\forall x \in X)$, 则称 x^0 是 ϕ (在 X 上) 的最小点.

定理 8.1.8 设 $x^0 \in X$, $y^0 \in \phi(x^0) \in \mathscr{D}$, K^* 是 K 的对偶锥.

(1) x^0 是 ϕ 的最小点当且仅当

$$\phi^T(x^0, y^0)(u) \subset K \quad \forall u \in X. \quad (8.1.3)$$

(2) x^0 是 ϕ 的最小点当且仅当

$$0 \in \partial^T \phi(x^0, y^0)(q^*) \quad \forall q^* \in K^*. \quad (8.1.4)$$

证明 (1) 设 (8.1.3) 成立, 则对任意的 $x \in X$, 由定理 8.1.7 得

$$\phi(x) \subset y^0 + \phi^T(x^0, y^0)(x - x^0) \subset y^0 + K.$$

因此, 由定义 8.1.2 可知, x^0 是 ϕ 的最小点.

反之, 设 x^0 是 ϕ 的最小点, 则对任意的 $v \in \phi^T(x^0, y^0)(u)$ 和任意的 $\varepsilon > 0$, 存在 $\bar{u} \in u + \varepsilon B$ ($B \subset \mathscr{D}$ 是单位球) 和 $t > 0$, 使得

$$v \in \frac{\phi(x^0 + t\bar{u}) - y^0}{t}.$$

由于 K 是闭的, 令 $\varepsilon \rightarrow 0$ 得 $v \in K$, 即 (8.1.3) 成立.

(2) 我们证明 (8.1.3) 和 (8.1.4) 等价. 事实上, 对任意的 $u \in \text{dom } \phi^T(x^0, y^0)$, $v \in \phi^T(x^0, y^0)(u) \subset K$, $q^* \in K^*$, 有

$$\langle 0, u \rangle - \langle q^*, v \rangle \geq 0.$$

由定理 8.1.1 即得 $0 \in \partial^T \phi(x^0, y^0)(q^*)$. 反之, 设 $0 \in \partial^T \phi(x^0, y^0)(q^*) (\forall q^* \in K^*)$, 对于 $v \in \phi^T(x^0, y^0)(u)$, 则有

$$\langle 0, u \rangle - \langle q^*, v \rangle \leq 0, \quad \forall q^* \in K^*,$$

故 $v \in K$. 于是, 定理得证. \square

定理 8.1.9 设 $S \subset \mathscr{B}$ 和 $M \subset \mathscr{D}$ 是非空闭凸集, $\mathscr{A} \in$

$L(\mathcal{B}, \mathcal{D})$ 是线性算子. 定义集值映射 $\phi: \mathcal{B} \rightarrow 2^{\mathcal{D}}$,

$$\phi(x) = \begin{cases} \mathcal{A}(x) - M, & x \in S; \\ \emptyset, & x \notin S, \end{cases}$$

则对任意的 $x \in S$ 和 $y \in \mathcal{A}(x) - M$, 有

$$\phi^T(x, y)(u) = \begin{cases} \mathcal{A}(x) - T_M(\mathcal{A}(x) - y), & u \in T_S(x); \\ \emptyset, & u \notin T_S(x) \end{cases}$$

和

$$\partial^T \phi(x, y)(q^*) = \begin{cases} \mathcal{A}^*(q^*) - N_S(x), & q \in N_M(\mathcal{A}(x) - y); \\ \emptyset, & q \notin N_M(\mathcal{A}(x) - y). \end{cases}$$

证明 设 $v \in \phi^T(x, y)(u)$, 由定义 8.1.1 知存在序列 $\{u^k\}$, $\{v^k\}$ 和 $\{t_k\} (t_k > 0, k = 1, 2, \dots)$, $u^k \rightarrow u$, $v^k \rightarrow v$, $t_k \rightarrow 0$, 使得

$$y + t_k v^k \in \phi(x + t_k u^k), \quad k = 1, 2, \dots.$$

也就是说有 $x + t_k u^k \in S (k = 1, 2, \dots)$, 即 $u \in T_S(x)$, 并且有

$$\mathcal{A}(x) - y + t_k (\mathcal{A}(u^k) - v^k) \in M, \quad k = 1, 2, \dots.$$

因此, 我们有 $\mathcal{A}(u) - v \in T_M(\mathcal{A}(x) - y)$.

反之, 设 $u \in T_S(x)$ 和 $v \in \mathcal{A}(u) - T_M(\mathcal{A}(x) - y)$, 则存在 $u^k \rightarrow u$, $w^k \rightarrow \mathcal{A}(u) - v$ 和 $t_k^1, t_k^2 > 0$, 使得

$$x + t_k^1 u^k \in S, \quad k = 1, 2, \dots,$$

$$\mathcal{A}(x) - y + t_k^2 w^k \in M, \quad k = 1, 2, \dots.$$

设 $t^k = \min(t_k^1, t_k^2)$ 和 $v^k = \mathcal{A}(u^k) - w^k$, 则有 $v^k \rightarrow v$. 从上式得

$$y + t_k v^k \in \phi(x + t_k u^k), \quad k = 1, 2, \dots,$$

故有 $v \in \phi^T(x, y)(u)$. 由定义 8.1.1 可知, $p^* \in \phi^T(x, y)(q^*)$ 当且仅当

$$\sup_{u \in T_S(x)} \sup_{w \in T_M(\mathcal{A}(x) - y)} [\langle p^*, u \rangle + \langle -q^*, \mathcal{A}(u) - w \rangle]$$

$$= \sup_{u \in T_S(p^*)} \langle p^* - \mathcal{A}^*(q^*), u \rangle + \sup_{w \in T_M(\mathcal{A}(x) - y)} \langle q^*, w \rangle \leq 0.$$

此即 $p^* - \mathcal{A}^*(q^*) \in N_S(x)$ 和 $q^* \in N_M(\mathcal{A}(x) - y)$, 从而定理得证. \square

下面研究集值映射的切导数和切微分的计算问题.

再设 \mathcal{B}_1 是 Banach 空间, $\psi: \mathcal{B} \rightarrow 2^{\mathcal{B}_1}$, $x \mapsto \psi(x)$ 是集值映射.

定义 8.1.3 若集值映射 ψ 的图象 $\text{graph } \psi$ 是闭的或凸的, 则分别称 ψ 在 \mathcal{B} 上是闭的或凸的.

定理 8.1.10 设集值映射 ψ 在 \mathcal{B} 上是闭凸的, $\mathcal{A} \in L(\mathcal{B}_1, \mathcal{B})$. 若

$$0 \in \text{int}(\text{Im } \mathcal{A} - \text{dom } \psi),$$

则有

$$(\psi \circ \mathcal{A})^T(z, y) = \psi^T(\mathcal{A}(z), y) \mathcal{A}, \quad (8.1.5)$$

$$\partial^T(\psi \circ \mathcal{A})(z, y) = \mathcal{A}^* \partial^T \psi(\mathcal{A}(z), y). \quad (8.1.6)$$

证明 作集值映射 $g: \mathcal{B}_1 \rightarrow 2^{\mathcal{B}}$, $z \mapsto g(z) = \psi \circ \mathcal{A}(z)$. 由于 \mathcal{A} 是线性算子, ψ 在 \mathcal{B} 上是闭凸的, 因此 g 在 \mathcal{B}_1 上是闭凸的. $\text{graph } g = (\mathcal{A} \times 1)^{-1} \text{graph } \psi$, 于是根据定理 5.1.13, 我们有

$$T_{\text{graph } g}(z, y) = (\mathcal{A} \times 1)^{-1} T_{\text{graph } \psi}(\mathcal{A}(z), y),$$

$$N_{\text{graph } g}(z, y) = (\mathcal{A}^* \times 1) N_{\text{graph } \psi}(\mathcal{A}(z), y).$$

从上式即可推得 (8.1.5) 和 (8.1.6). \square

定理 8.1.11 设集值映射 ψ 在 \mathcal{B} 上是凸的, $\mathcal{A} \in L(\mathcal{B}, \mathcal{B}_1)$, 作集值映射 $\mathcal{A}\psi: \mathcal{B} \rightarrow 2^{\mathcal{B}_1}$, $x \mapsto \mathcal{A}\psi(x) = \bigcup_{y \in \psi(x)} \mathcal{A}(y)$. 若 $\mathcal{A}\psi$ 在 \mathcal{B} 上是凸的, 并且 $0 \in \text{int}(\text{Im } \mathcal{A} - \text{dom } \psi)$, 则

$$(\mathcal{A}\psi)^T(x, \mathcal{A}(y)) = \text{cl}(\mathcal{A}\psi^T(x, y)), \quad (8.1.7)$$

$$\partial^T(\mathcal{A}\psi)(x, \mathcal{A}(y)) = \partial^T \psi(x, y) \mathcal{A}^*. \quad (8.1.8)$$

证明 由 $\text{graph}(\mathcal{A}\psi) = (1 \times \mathcal{A})\text{graph } \psi$ 成立. 于是, 由定

理 5.1.11 得

$$T_{\text{graph}(\mathcal{A}\phi)}(x, \mathcal{A}(y)) = \text{cl}\{(1 \times \mathcal{A})T_{\text{graph}\phi}(x, y)\},$$

$$N_{\text{graph}(\mathcal{A}\phi)}(x, \mathcal{A}(y)) = \{(1 \times \mathcal{A})^*{}^{-1}N_{\text{graph}\phi}(x, y)\}.$$

从上面两式, 即可得知 (8.1.7) 和 (8.1.8) 成立. 定理得证. \square

再设 \mathcal{B}_2 是 Banach 空间.

定理 8.1.12 设集值映射 ϕ 在 \mathcal{B} 上是闭凸的, $\mathcal{A} \in L(\mathcal{B}_2, \mathcal{B})$, $\mathcal{A}_1 \in L(\mathcal{D}, \mathcal{B}_1)$, 集值映射 $\mathcal{A}_1\phi: \mathcal{B} \rightarrow 2^{\mathcal{B}_1}$ 是凸的, 又 $u \in \mathcal{B}_2$, $y \in \phi(\mathcal{A}(u))$. 若 $\theta \in \text{int}(\text{Im}\mathcal{A} - \text{dom}\phi)$, 则

$$(\mathcal{A}_1\phi\mathcal{A})^T(u, \mathcal{A}_1(y)) = \text{cl}\{\mathcal{A}_1\phi^T(\mathcal{A}(u), y)\mathcal{A}\},$$

$$\partial^T(\mathcal{A}_1\phi\mathcal{A})(u, \mathcal{A}_1(y)) = \mathcal{A}^*\partial^T\phi(\mathcal{A}(u), y)\mathcal{A}_1^*.$$

证明 由定理 8.1.10 和定理 8.1.11 可得证. \square

定理 8.1.13 设 $\phi_1: \mathcal{B} \rightarrow 2^{\mathcal{B}_1}$ 和 $\phi_2: \mathcal{D} \rightarrow 2^{\mathcal{B}_2}$ 是集值映射, $\mathcal{A} \in L(\mathcal{B}, \mathcal{D})$, $\text{graph}\phi_1$ 和 $\text{graph}\phi_2$ 是闭凸的, 并且 $\theta \in \text{int}(\mathcal{A}\text{dom}\phi_1 - \text{dom}\phi_2)$. 若 $x \in \text{dom}\phi_1 \cap \mathcal{A}^{-1}\text{dom}\phi_2$, $y \in \phi_1(x)$, $z \in \phi_2(\mathcal{A}(x))$, 则

$$(\phi_1 + \phi_2\mathcal{A})^T(x, y+z) = \text{cl}\{\phi_1^T(x, y) + \phi_2^T(\mathcal{A}(x), z)\mathcal{A}\},$$

$$\partial^T(\phi_1 + \phi_2\mathcal{A})(x, y+z) = \partial^T\phi_1(x, y) + \mathcal{A}^*\partial^T\phi_2(\mathcal{A}(x), z).$$

证明 集值映射 $\phi_1 + \phi_2\mathcal{A}: \mathcal{B} \rightarrow 2^{\mathcal{B}_1}$ 可以分解成以下三个映射的复合:

$$\mathcal{A}_1: x \in \mathcal{B} \rightarrow (x, \mathcal{A}(x)) \in \mathcal{B} \times \mathcal{D},$$

$$\phi_3: (x, y) \in \mathcal{B} \times \mathcal{D} \rightarrow \phi_1(x) \times \phi_2(y) \subset \mathcal{B}_1 \times \mathcal{B}_2,$$

$$A_2: (z_1, z_2) \in \mathcal{B}_1 \times \mathcal{B}_2 \rightarrow z_1 + z_2 \in \mathcal{B}_1.$$

也即 $\phi_1 + \phi_2\mathcal{A} = \mathcal{A}_2\phi_3\mathcal{A}_1$, 并且有

$$\theta \in \text{int}(\text{Im } \mathcal{A}_1 - \text{dom } \psi_2).$$

设 $\phi_3^T(x, y, z_1, z_2) = \phi_1^T(x, z_1) \times \phi_2^T(y, z_2)$, 由定理 8.1.12 即得结论. \square

推论 8.1.14 设 $\text{graph } \phi_1$ 和 $\text{graph } \phi_2$ 是凸的, $\theta \in \text{int}(\text{dom } \phi_1 - \text{dom } \phi_2)$, $x \in \text{dom } \phi_1 \cap \text{dom } \phi_2$, $y \in \phi_1(x)$, $z \in \phi_2(x)$, 则

$$(\phi_1 + \phi_2)^T(x, y + z) = \text{cl}\{\phi_1^T(x, y) + \phi_2^T(x, z)\},$$

$$\partial^T(\phi_1 + \phi_2)(x, y + z) = \partial^T\phi_1(x, y) + \partial^T\phi_2(x, z).$$

证明 由定理 8.1.13 可得证. \square

推论 8.1.15 设集值映射 ϕ 在 \mathcal{B} 上是闭凸的, $S \subset \mathcal{B}$ 是闭凸集, 集值映射 $\theta_S: S \rightarrow 2^{\mathcal{C}}$,

$$\theta_S(x) = \begin{cases} 0, & x \in S; \\ \emptyset, & x \notin S. \end{cases}$$

记 $\phi|_S = \phi + \theta_S$, 若 $\theta \in \text{int}(\text{dom } \phi - S)$, 则对任意的 $x \in S$ 和 $y \in \phi(x)$, 有

$$(\phi|_S)^T(x, y) = \phi^T(x, y) + \theta_S(x),$$

$$\partial^T(\phi|_S)(x, y)(q) = \partial^T\phi(x, y)(q) + N_S(x).$$

证明 在推论 8.1.14 中令 $\phi_1 = \phi$, $\phi_2 = \theta_S$, 即可得证. \square

推论 8.1.16 设集值映射 ϕ 在 \mathcal{B} 上是闭凸的, $K \subset \mathcal{C}$ 是闭凸锥, 记集合 $\phi_+(x) = \phi(x) + K (\forall x \in \mathcal{B})$, 则对任意 $x \in \text{dom } \phi$ 和任意 $y \in \phi(x)$, 有

$$\phi_+^T(x, y)(u) = \phi^T(x, y)(u) + K,$$

$$\partial^T\phi_+(x, y)(q^*) = \begin{cases} \partial^T\phi(x, y)(q^*), & q^* \in K^*; \\ \emptyset, & q^* \notin K^*. \end{cases}$$

证明 在推论 8.1.14 中令 $\phi_1 = \phi$, $\phi_2 = K$, 即可得证. \square

定理 8.1.17 设集值映射 $\phi_i: \mathcal{B} \rightarrow 2^{\mathcal{D}} (i = 1, \dots, m)$ 是闭凸的. 若有 $x^0 \in \bigcap_{i=1}^m \text{int}(\text{dom} \phi_i)$, 并且存在 $\gamma > 0$, 使得

$$\bigcap_{i=1}^m (\phi_i(x^0 + u_i) - v_i) \neq \emptyset \quad \forall (u_i, v_i) \in \gamma(B \times B),$$

则对任意的 $x \in \bigcap_{i=1}^m \text{dom} \phi_i$ 和任意的 $y \in \bigcap_{i=1}^m \phi_i(x)$, 有

$$\left(\bigcap_{i=1}^m \phi_i \right)^T(x, y) = \bigcap_{i=1}^m \phi_i^T(x, y).$$

证明 由于 $\text{graph} \left(\bigcap_{i=1}^m \phi_i \right) = \bigcap_{i=1}^m \text{graph} \phi_i$, 由已设可知存在 $\delta > 0$, 使得

$$\bigcap_{i=1}^m (\text{graph} \phi_i - (u_i, v_i)) \neq \emptyset \quad \forall (u_i, v_i) \in \delta(B \times B).$$

事实上, 取 $x^0 \in \bigcap_{i=1}^m \text{dom} \phi_i$ 和 $\delta \leq \gamma$, 使得

$$x^0 + \delta B \subset \bigcap_{i=1}^m \text{dom} \phi_i,$$

选择 $y^0 \in \bigcap_{i=1}^m (\phi_i(x^0 + u_i) - v_i)$, 则由已知得

$$(x^0, y^0) \in \bigcap_{i=1}^m (\text{graph} \phi_i - (u_i, v_i)).$$

于是, 由定理 5.1.17 即得定理的结论. \square

§ 8.2 D -切导数

现在, 利用 D -切锥(相依切锥)来定义一般集值映射的 D -切导数.

设 \mathcal{B} 和 \mathcal{D} 是 Banach 空间, $\phi: \mathcal{B} \rightarrow 2^{\mathcal{D}}, x \mapsto \phi(x)$ 是集值映射, 点 $x^0 \in \mathcal{B}, y^0 \in \phi(x^0), T_{\text{graph} \phi}^D(x^0, y^0)$ 是 ϕ 的图象 $\text{graph} \phi$ 在点 (x^0, y^0) 处的 D -切锥.

定义 8.2.1 设点 $x^0 \in \mathcal{B}$, $y^0 \in \psi(x^0)$, $u \in \mathcal{B}$, $\psi^D(x, y): \mathcal{B} \rightarrow 2^{\mathcal{D}}$, $u \mapsto \psi^D(x^0, y^0)(u)$ 是集值映射. 若 $v \in \psi^D(x^0, y^0)(u)$ 当且仅当 $(u, v) \in T_{\text{graph } \psi}^D(x^0, y^0)$, 则称 $\psi^D(x^0, y^0)$ 是集值映射 ψ 在点 (x^0, y^0) 处的 D -切导数.

注 8.2.1 由注 5.2.1 我们知道, $v^0 \in \psi^D(x^0, y^0)(u^0)$ 当且仅当存在 $t_k \rightarrow 0^+$, $u^k \rightarrow u^0$ 和 $v^k \rightarrow v^0$, 使得 $v^k \in (\psi(x^0 + t_k u^k) - y^0)/t_k$ ($k = 1, 2, \dots$), 或者

$$\liminf_{\substack{u \rightarrow u^0 \\ t \rightarrow 0^+}} d\left(v^0, \frac{\psi(x^0 + tu) - y^0}{t}\right) = 0. \quad (8.2.1)$$

特别是, 当 ψ 是单值映射时, 上式可写作

$$\liminf_{\substack{u \rightarrow u^0 \\ t \rightarrow 0^+}} \frac{\|\psi(x^0 + tu) - \psi(x^0) - t v^0\|}{t} = 0.$$

进一步, 若单值映射 ψ 在点 x^0 处是严格可微的, 则

$$\psi^D(x^0, y^0)(u^0) = D_S \psi(x^0) u^0.$$

当 $\text{graph } \psi$ 是凸的, 则有 $\psi^D(x^0, y^0)(u^0) = \psi^D(x^0, y^0)(u^0)$.

注 8.2.2 集值映射 ψ 在点 x^0 处是 Lipschitz 的, 即存在常数 $\gamma > 0$ 使得

$$\psi(x^0 + tu) - y^0 \subset \psi(x^0 + tu^0) - y^0 + \gamma t \|u - u^0\| B.$$

定理 8.2.1 设 $x \in \text{int}(\text{dom } \psi)$. 若集值映射 ψ 在 x 附近是 Lipschitz 的, 则 $v \in \psi^D(x, y)(u)$ 当且仅当

$$\liminf_{t \rightarrow 0^+} d\left(v, \frac{\psi(x + tu) - y}{t}\right) = 0. \quad (8.2.2)$$

当 \mathcal{D} 是有限维空间时, 则有 $\text{dom}(\psi^D(x, y)) = \mathcal{B}$.

证明 必要性. 由于 ψ 在 x 处是 Lipschitz 的, 则有

$$\psi(x + tu') - y \subset \psi(x + tu) - y + \gamma t \|u' - u\| B.$$

令 $t \rightarrow 0^+$, $u^t \rightarrow u$ 由 (8.1.1) 即得 (8.2.2).

充分性. 设取任意的 $u \in \mathcal{B}$, 则由 (8.2.2) 对充分小的 t , 有

$$y \in \phi(x) \subset \phi(x + tu) + \gamma t \|u\| B.$$

故存在 $v' \in \phi(x + tu)$ 使得 $\frac{(v' - y)}{t} \in \gamma \|u\| B$. 由于 $\gamma \|u\| B$ 是紧的, 故存在收敛子列 $\{(v'_k - y)/t_k\}$, $(v'_k - y)/t_k \rightarrow v$ ($k \rightarrow \infty$), 由定义 8.2.1 便得 $v \in \phi^D(x, y)(u)$. 当 ϕ 是有限维空间时, 由上述结果易知第 2 个结论成立. \square

定理 8.2.2 设 $X \subset \mathcal{B}$ 是非空开集. 若映射 $\varphi: X \rightarrow \mathcal{L}$ 是严格可微的, 集合 $S \subset X$ 非空, $x \in S$, 则

$$\varphi|_S^D(x)(u) = \begin{cases} D_S \varphi(x)u, & u \in T^D(x); \\ \emptyset, & u \notin T^D(x). \end{cases}$$

证明 当 $u \in T_S^D(x)$ 时, 则存在 $t_k \rightarrow 0^+$, $u^k \rightarrow u$ ($k \rightarrow \infty$), 使得 $x + t_k u^k \in S$. 因为

$$\begin{aligned} \varphi|_S(x + t_k u^k) &= \varphi(x + t_k u^k) \\ &= \varphi(x) + t_k (D_S \varphi(x) u^k + o(t_k)), \end{aligned}$$

令 $v^k = D_S \varphi(x) u^k + o(t_k)$, 则 $v^k \rightarrow D_S \varphi(x) u$, 并且有

$$v^k \in \{\varphi|_S(x + t_k u^k) - \varphi|_S(x)\} / t_k.$$

由此, 得到 $\varphi^D(x, \varphi(x))(u) = D_S \varphi(x)u$. \square

设 \mathcal{B}_1 是 Banach 空间.

定理 8.2.3 设 $x \in \mathcal{B}$, $y \in \phi(x)$, Ω 是 $\text{Im } \phi$ 的一个开邻域. 若映射 $\varphi: \Omega \rightarrow \mathcal{B}_1$ 是严格可微的, 则

$$\begin{aligned} D_S \varphi(y) \phi^D(x, y)(u) &\subset (\varphi \phi)^D(x, \varphi(y))(u) \\ &\quad \forall u \in \mathcal{B}. \end{aligned} \tag{8.2.3}$$

再若 ϕ 在 $x \in \text{int}(\text{dom } \phi)$ 处是 Lipschitz 的, 并且是紧值的, \mathcal{B}_1 是有限维空间, 则 (8.2.3) 成为等式.

证明 作映射 $(1 \times \varphi): (x, y) \in \mathcal{B} \times \Omega \mapsto (x, \varphi(y)) \in \mathcal{D} \times \mathcal{B}_1$ 对集值映射 $\theta = \varphi\phi: \mathcal{B} \rightarrow 2^{\mathcal{D}}$, $x \mapsto \varphi\phi(x) \subset \mathcal{B}_1$ 有

$$\text{graph } \theta = (1 \times \varphi)\text{graph } \phi.$$

据此,由定理 4.2.6 即可得证.

为证明 (8.2.3) 成为等式, 设 $w \in (\varphi\phi)^D(x, \varphi(y))(u)$, 则存在 $t_k \rightarrow 0$, $u^k \rightarrow u$ 和 $w^k \rightarrow w$ ($k \rightarrow \infty$), 使得

$$t_k w^k \in \varphi[\phi(x + t_k u^k) - \varphi(y)], k = 1, 2, \dots.$$

因此, 存在 v^k ($k = 1, 2, \dots$) $\in [\phi(x + t_k u^k) - y]/t_k$, 使得

$$t_k w^k = \varphi(y + t_k v^k) - \varphi(y), k = 1, 2, \dots.$$

因为 ϕ 在 x 处是 Lipschitz 的, 故有 $v^k \in \phi(x) - y + t^k \|u^k\| B$. 再由 \mathcal{D} 是有限维空间和 ϕ 是紧值的, 所以 $\{v^k\}$ 中存在收敛子列, 不妨设为 $v^k \rightarrow v$. 于是有 $v \in \phi^D(x, y)(u)$, 并且 $w^k \rightarrow D_S \varphi(y) v = w$. 由此得到 $(\varphi\phi)^D(x, \varphi(y))(u) \subset D_S \varphi(y) \phi^D(x, y)(u)$. \square

定理 8.2.4 设 $\phi_i: \mathcal{B} \rightarrow 2^{\mathcal{D}}$, $x \mapsto \phi_i(x)$ ($i = 1, 2, \dots, m$) 是集值映射, $x \in \text{dom } \phi_i$ ($i = 1, 2, \dots, m$). 若 $y \in \bigcap_{i=1}^m \text{Im } \phi_i(x)$, 则

$$\left(\bigcup_{i=1}^m \phi_i \right)^D(x, y)(u) = \bigcup_{i=1}^m \phi_i^D(x, y)(u).$$

证明 根据定理 4.2.5 和定理 4.2.7, 可以得证. \square

§ 8.3 C-切导数和 C-切微分

本节将用 C-切锥, 引进一般集值映射的 C-切导数和 C-切微分.

设 \mathcal{B} 和 \mathcal{D} 是 Banach 空间, \mathcal{B}^* 和 \mathcal{D}^* 分别是 \mathcal{B} 和 \mathcal{D} 的对偶空间, $\phi: \mathcal{B} \rightarrow 2^{\mathcal{D}}$, $x \mapsto \phi(x)$ 是集值映射, $(x^0, y^0) \in \text{graph } \phi$, $T_{\text{graph } \phi}^C(x^0, y^0)$ 是 $\text{graph } \phi$ 在点 (x^0, y^0) 处的 C-切锥.

定义 8.3.1 设点 $x^0 \in \mathcal{B}$, $y^0 \in \phi(x^0)$, $u \in \mathcal{B}$, 考虑集值映

射 $\psi^c(x^0, y^0): \mathcal{B} \rightarrow 2^{\mathcal{V}}$, $u \mapsto \psi^c(x^0, y^0)(u)$. 若 $v \in \psi^c(x^0, y^0)(u)$ 当且仅当

$$(u, v) \in T_{\text{graph} \psi}(x^0, y^0),$$

则称 $\psi^c(x^0, y^0)$ 是集值映射 ψ 在点 (x^0, y^0) 处的 C -切导数. 考虑集值映射 $\partial^c \psi(x^0, y^0): \mathcal{B}^* \rightarrow 2^{\mathcal{V}^*}$, $q_0^* \mapsto \partial^c \psi(x^0, y^0)(q_0^*)$. 若 $p_0^* \in \partial^c \psi(x^0, y^0)(q_0^*)$ 当且仅当

$$\langle p_0^*, u \rangle \leq \langle q_0^*, v \rangle \quad \forall u \in \mathcal{B}, \forall v \in \psi^c(x^0, y^0)(u),$$

则称 $\partial^c \psi(x^0, y^0)$ 是集值映射 ψ 在点 (x^0, y^0) 处的 C -切微分.

由定义 5.2.2 可以推得以下性质.

定理 8.3.1 设 $x^0 \in \mathcal{B}$, $y^0 \in \psi(x^0)$, 则 $v^0 \in \psi^c(x^0, y^0)(u^0)$ 与下面两命题等价:

(1) 若对任意的 $\epsilon_1, \epsilon_2 > 0$, 存在 $\alpha, \beta > 0$, 使得 $(x, y) \in B_{\text{graph} \psi}(x^0, y^0, \alpha)$, 则对任意的 $t(\alpha, \beta)$, 存在 $u \in u^0 + \epsilon_1 B$, $v \in v^0 + \epsilon_2 B$, 使得

$$v \in (\psi(x + tu) - y)/t.$$

(2) 对所有的收敛于 $(x^0, y^0, 0)$ 的序列 $(x^k, y^k, t_k) \in \text{graph} \psi \times (0, +\infty)$, 存在序列 $\{u^k\}$ 和 $\{v^k\}$, $u^k \rightarrow u^0$, $v^k \rightarrow v^0$ ($k \rightarrow \infty$), 使得

$$y^k + t_k v^k \in \psi(x^k + t_k u^k), \quad k = 1, 2, \dots.$$

定理 8.3.2 设 $x^0 \in \mathcal{B}$, $y^0 \in \psi(x^0)$. 若 ψ 在 $x^0 \in \text{int}(\text{dom} \psi)$ 处是 Lipschitz 的, 则 $v^0 \in \psi^c(x^0, y^0)(u^0)$ 当且仅当

$$\lim_{\substack{x \rightarrow x^0 \\ t \rightarrow 0^+}} d\left(v^0, \frac{\psi(x + tu^0) - y^0}{t}\right) = 0.$$

定理 8.3.3 设 $x^0 \in R^n$, f 是单值连续可微函数, $y^0 = f(x^0)$, 则

$$f^c(x^0, f(x^0)) = \nabla f(x^0).$$

由定义 7.3.1, 定义 7.2.1 以及定理 7.2.3, 可以得到 C-弱方向导数和 C-弱次微分与 C-切导数和 C-切微分之间的关系.

定理 8.3.4 设 $x^0 \in \mathcal{B}$, $y^0 \in \psi(x^0)$, $K \subset \mathcal{B}$ 是内部非空的尖闭凸锥, $p \in \text{int}K$, $d \in \mathcal{B}$, $\Psi(x^0, y^0, d)$ 由 (7.2.1) 确定, $\hat{D} = \mathcal{B}$. 又设 $\Psi(x^0, y^0, d)$ 是 K -有界的.

$$(1) \phi^c(x^0, y^0)(d) \subset \phi_\omega^c(x^0, y^0; d).$$

$$(2) \text{ 若 } x^* \in \partial^c \phi(x^0, y^0)(q^*), q^* \in K^* \setminus \{0\}, \text{ 则}$$

$$x^* \left(\langle q^*, p \rangle \right)^{-1} \in \partial_\omega^c \phi(x^0, y^0)_p.$$

$$(3) \text{ 若 } x^* \in \partial_\omega^c \phi(x^0, y^0)_p, \text{ 则存在 } q^* \in K^* \setminus \{0\} \text{ 使得}$$

$$\langle q^*, p \rangle x^* \in \partial^c \phi(x^0, y^0)(q^*).$$

定理 8.3.5 设 $\Omega \subset \mathcal{B}$ 是非空开集. 若函数 $f: \Omega \rightarrow R$ 在 $x \in \Omega$ 处是连续可微的, 集合 $S \subset \mathcal{B}$, $x \in S$, 则

$$f|_S(x, f(x))(u) = \begin{cases} \nabla f(x)u, & u \in T_S(x); \\ \emptyset, & u \notin T_S(x). \end{cases}$$

证明 设序列 $\{x^k, t_k\} \in S \times (0, 1)$ 收敛于 $(x, 0)$, 当 $u \in T_S(x)$ 时, 则存在序列 $\{u^k\}$, $u^k \rightarrow u (k \rightarrow \infty)$, 使得

$$x^k + t_k u^k \in S, k = 1, 2, \dots,$$

有

$$\begin{aligned} f|_S(x^k + t_k u^k) &= f(x^k + t_k u^k) \\ &= f(x^k) + t_k (\nabla f(x^k)u^k + O(t_k)). \end{aligned}$$

因为 f 是连续可微的, 所以序列 $\{v^k\} = \{\nabla f(x^k)u^k + O(t_k)\}$ 收敛于 $\nabla f(x)u$, 由此有

$$f|_S(x^k) + t_k v^k = f|_S(x + t_k u^k), k = 1, 2, \dots.$$

根据定理 8.3.1 即知结论成立. \square

定理 8.3.6 设 $\Omega \subset \mathcal{B}$ 是开集, $\mathcal{A}: \Omega \rightarrow \mathcal{B}$ 是连续可微算子,

$S \subset \mathcal{B}$ 和 $Y \subset \mathcal{D}$ 分别是 \mathcal{B} 和 \mathcal{D} 中的闭集. 若集值映射 $\phi: \mathcal{B} \rightarrow 2^{\mathcal{D}}$,

$$\phi(x) = \begin{cases} \mathcal{A}(x) - Y, & x \in S; \\ \emptyset, & x \notin S, \end{cases}$$

并且 $(x, y) \in \text{graph } \phi$, 则 $v \in \phi^c(x, y)(u)$ 当且仅当 $u \in T_S(x)$ 和 $v \in \mathcal{A}'_x(u) - T_Y(\mathcal{A}(x) - y)$.

证明 必要性. 设 $v \in \phi^c(x, y)(u)$, 取收敛于 $(x, \mathcal{A}(x) - y, 0)$ 的序列 $\{x^k, z^k, t_k\} \in S \times Y \times (0, +\infty)$, 则序列 $y^k = \mathcal{A}(x^k) - z^k$ 收敛于 y . 由定义 8.3.1 可知, 存在序列 $\{u^k\}$ 和 $\{v^k\}$, $u^k \rightarrow u$, $v^k \rightarrow v$ ($k \rightarrow \infty$), 使得

$$x^k + t_k u^k \in S, \mathcal{A}(x^k + t_k u^k) \in Y + y^k + t_k v^k, k = 1, 2, \dots.$$

这意味着有 $u \in T_S(x)$ 和

$$w^k = \mathcal{A}(x^k + t_k u^k) - \mathcal{A}(x^k) - v^k \rightarrow \mathcal{A}'_x(u) - v,$$

$$z^k + t_k w^k \in Y, k = 1, 2, \dots.$$

由此, 得 $v \in \mathcal{A}'_x(u) - T_Y(\mathcal{A}(x) - y)$.

充分性. 取收敛于 $(x, y, 0)$ 的序列 $\{x^k, y^k, t_k\} \in \text{graph } \phi \times (0, +\infty)$. 由 $u \in T_S(x)$ 和 $v \in \mathcal{A}'_x(u) - T_Y(\mathcal{A}(x) - y)$ 得知存在序列 $\{u^k\}$, $u^k \rightarrow u$ ($k \rightarrow \infty$), 使得

$$x^k + t_k u^k \in S, k = 1, 2, \dots.$$

于是, 存在序列 $\{w^k\}$, $w^k \rightarrow \mathcal{A}'_x(u) - v$ ($k \rightarrow \infty$), 使得

$$\mathcal{A}(x^k) - y^k + t_k w^k \in Y, k = 1, 2, \dots.$$

据此, 则有 $v^k = \{[\mathcal{A}(x^k + t_k u^k) - \mathcal{A}(x^k)]/t_k\} - w^k \rightarrow v^0$ ($k \rightarrow \infty$), 使得 $y^k + t_k v^k \in \phi(x^k + t_k u^k)$ ($k = 1, 2, \dots$). \square

定理 8.3.7 设 \mathcal{B} 是 Hilbert 空间, $S \subset \mathcal{B}$ 是闭凸集, $N_S(x)$ 是 S 在点 x 处的法锥, $p \in N_S(x)$, 作集值映射 $N_S: \mathcal{B} \rightarrow 2^{\mathcal{B}^*}$, $x \mapsto N_S(x)$. 若单值映射 $\pi_S: x \rightarrow \pi_S(x) \in S$, $\pi_S(x)$ 使得下式成立:

$$\pi_S^{-1}(x) = x + N_S(x),$$

则 $q \in N_S^C(x, p)(u)$ 当且仅当 $u \in \pi_S^C(x, p)(u + q)$.

证明 显然, $p \in N_S(x)$ 当且仅当 $x = \pi_S(x + p)$.

必要性. 设 $q \in N_S^C(x, p)(u)$, 考虑序列 $\{(y^k, t_k)\} \subset \mathcal{B} \times (0, +\infty)$,

$$(y^k, t_k) \rightarrow (x + p, 0), \quad k \rightarrow \infty.$$

作序列 $\{x^k\} = \{\pi_S(y^k)\}$, 则 $x^k \rightarrow x = \pi_S(x + p)$, $p^k = y^k - x^k \rightarrow p (k \rightarrow \infty)$, 于是存在序列 $\{u^k\}$ 和 $\{q^k\}$, $u^k \rightarrow u$, $q^k \rightarrow q (k \rightarrow \infty)$, 使得

$$p^k + t_k q^k \in N_S(x^k + t_k u^k), \quad k = 1, 2, \dots,$$

即

$$\pi_S(y^k) + t_k u^k = \pi_S(y^k + t_k(q^k + u^k)), \quad k = 1, 2, \dots.$$

因此, 我们有 $u \in \pi_S^C(x, p)(u + q)$.

充分性. 设 $u \in \pi_S^C(x, p)(u + q)$, 序列 $\{x^k, p^k, t_k\} \in \text{graph } N_S \times (0, +\infty)$ 收敛于 $(x, p, 0)$. 因为 $x^k + p^k$ 收敛于 $x + p$, 所以存在序列 $\{u^k\}$ 和 $\{w^k\}$, $u^k \rightarrow u$, $w^k \rightarrow u + q (k \rightarrow \infty)$, 使得

$$\begin{aligned} x^k + t_k u^k &= \pi_S(x^k + p^k) + t_k u^k \\ &= \pi_S(x^k + p^k + t_k w^k), \quad k = 1, 2, \dots. \end{aligned}$$

令 $q^k = w^k - u^k$, 则有 $q^k \rightarrow q$. 从上式得 $p^k + t_k q^k \in N_S(x^k + t_k u^k)$ ($k = 1, 2, \dots$), 故 $q \in N_S^C(x, p)(u)$. \square

定理 8.3.8 设 \mathcal{B}_0 , \mathcal{B} 和 \mathcal{D} 均为有限维 Euclid 空间, $x' \in \mathcal{B}$, $\phi(x') \subset \mathcal{D}$ 是闭的, $\mathcal{A} : \mathcal{B}_0 \rightarrow \mathcal{B}$ 是连续可微映射, $x \in \mathcal{A}^{-1} \text{dom } \phi$ 和 $y \in \phi(x)$. 若

$$\text{Im } J_{\mathcal{A}}(x) = \text{dom}[\phi'(\mathcal{A}(x), y)] = \mathcal{B},$$

则

$$(\phi \circ \mathcal{A})^C(x, y)(u) \supset \phi'(\mathcal{A}(x), y)(J_{\mathcal{A}}(x)u) \quad \forall u \in \mathcal{B},$$

$$\partial^c(\psi_{\mathcal{A}})(x, y)(q^*) \subset J_{\mathcal{A}}(x) \partial^c \psi(\mathcal{A}(x), y)(q^*)$$

$$\forall q^* \in \mathcal{D}^*.$$

证明 因为 $\text{graph} \psi_{\mathcal{A}} = (\mathcal{A} \times 1)^{-1} \text{graph} \psi$, 由推论 5.2.8 有

$$(J_{\mathcal{A}}(x) \times 1)^{-1} T_{\text{graph} \psi}^c(\mathcal{A}(x), y) \subset T_{\text{graph} \psi_{\mathcal{A}}}^c(x, y).$$

从假设知

$$\text{Im}(J_{\mathcal{A}}(x) \times 1) = T_{\text{graph} \psi}^c(\mathcal{A}(x), y) = \mathcal{B} \times \mathcal{D},$$

$$(\psi^c(\mathcal{A}(x), y) J_{\mathcal{A}}(x)) = J_{\mathcal{A}}(x) \partial^c \psi(\mathcal{A}(x), y).$$

由此, 可得定理结论. \square

定理 8.3.9 设集合 $S \subset \mathcal{B}$, $x \in S \cap \text{dom} \psi$, $y \in \psi(x)$, 则

$$(\psi|_S)^D(x, y)(u) \subset \psi^D(x, y) \Big|_{T_S^D(x)}(u). \quad (8.3.1)$$

若 \mathcal{B} 和 \mathcal{D} 是有限维空间, $\text{graph} \psi$ 和 S 都是闭的, 并且

$$T_S(x) = \text{dom}(\psi^c(x, y)) = \mathcal{B}, \quad (8.3.2)$$

则

$$\psi^c(x, y) \Big|_{T_S(x)}(u) \subset (\psi|_S)^c(x, y)(u),$$

$$\partial^c(\psi|_S)(x, y)(q^*) \subset \partial^c \psi(x, y)(q^*) + N_S(x) \quad \forall q \in \mathcal{D}^*.$$

证明 因为有 $\text{graph}(\psi|_S) = \text{graph}(\psi)(S \times \mathcal{D})$, 则

$$T_{\text{graph} \psi}^c|_S(x, y) \subset T_{\text{graph} \psi}^D(x, y) \cap (T_S^D(x) \times \mathcal{D}).$$

从上式即可得(8.3.1). 另外, 由(8.3.2)得知

$$T_{\text{graph} \psi}^c(x, y) = T_S(x) \times \mathcal{D} = \mathcal{B} \times \mathcal{D},$$

从而结论成立. \square

定理 8.3.10 设 \mathcal{B} 和 \mathcal{D} 是有限维空间, $\phi_i: \mathcal{B} \rightarrow 2^{\mathcal{D}} (i = 1, 2, \dots, m)$ 是集值映射, $(x, y) \in \bigcap_{i=1}^m \text{graph} \phi_i$. 若对任何 $(u_i, v_i) \in \mathcal{B} \times \mathcal{D} (i = 1, 2, \dots, m)$, 存在 $(u, v) \in \mathcal{B} \times \mathcal{D}$, 使得

$$v \in \phi_i^c(x, y)(u + u_i) - v_i, i = 1, 2, \dots, m,$$

则

$$\left(\bigcap_{i=1}^m \phi_i\right)^c(x, y)(u) \supset \bigcap_{i=1}^m \phi_i^c(x, y)(u) \quad \forall u \in \mathcal{B}.$$

证明 由推论 5.2.10 可使定理得证. \square

§ 8.4 M -次微分

最后, 介绍有限维空间中集值映射的另一种广义微分及其性质.

本节考虑集值映射 $\phi: R^n \rightarrow 2^{R^m}, x \mapsto \phi(x)$, 它的图象和核分别为

$$\text{graph} \phi = \{(x, y) \in R^n \times R^m \mid y \in \phi(x)\}$$

和

$$\text{Ker} \phi = \{x \in R^n \mid 0 \in \phi(x)\}.$$

设 $x^0 \in R^n$, 记集合

$$\limsup_{x \rightarrow x^0} \phi(x) = \{y \in R^m \mid \exists x^k \rightarrow x^0, y^k \rightarrow y,$$

$$k \rightarrow \infty; y^k \in \phi(x^k), k = 1, 2, \dots\},$$

并称它是集值映射 ϕ 在点 x^0 处的 KP -上极限集.

设集合 $S \subset R^n$ 非空, $x^0 \in S$, 记

$$P(x^0, S) = \{w \in \text{cl} S \mid \|x^0 - w\| = \inf_{x \in S} \|x - w\|\}.$$

定义 8.4.1 设集合 $S \subset R^n$ 非空, 点 $x^0 \in \text{cl} S$, 则称集合

$$N_S^M(x^0) = \lim_{x \rightarrow x^0} \sup [\text{cone}(x - P(x, S))]]$$

是集合 S 在点 x^0 处的 M -法锥. 若 $x^0 \notin \text{cl}S$, 则令 $N_S^M(x^0) = \emptyset$.

注 8.4.1 设 $x_1^0 \in S_1$, $x_2^0 \in S_2$, 由定义 8.4.1 我们有

$$N_{S_1 \times S_2}^M((x_1^0, x_2^0)) = N_{S_1}^M(x_1^0) \times N_{S_2}^M(x_2^0).$$

根据定义 5.2.2 的(2), 还有关系:

$$N_S^C(x^0) = \text{cl}(\text{co}N_S^M(x^0)).$$

定义 8.4.2 设点 $(x^0, y^0) \in \text{cl}(\text{graph}\phi)$, 定义集值映射 $\partial^M\phi(x^0, y^0): R^m \rightarrow 2^{R^n}$,

$$\partial^M\phi(x^0, y^0)(y^*) = \{x^* \in R^n \mid (x^*, -y^*) \in N_{\text{graph}\phi}^M(x^0, y^0)\}, \quad (8.4.1)$$

则称 $\partial^M\phi(x^0, y^0)$ 是集值映射 ϕ 在点 (x^0, y^0) 处的 M -次微分. 若 $(x^0, y^0) \notin \text{cl}(\text{graph}\phi)$, 令 $\partial^M\phi(x^0, y^0)(y^*) = \emptyset$, 若 ϕ 在点 x^0 处是单值的, 令 $\partial^M\phi(x^0) = \partial^M\phi(x^0, y^0)$. 若 $\partial^M\phi(x^0, y^0) \neq \emptyset$, 则称 ϕ 在点 (x^0, y^0) 处是 M -次可微的.

由定义 8.4.2 可知, $\partial^M\phi(x^0, y^0)$ 关于 y^* 是正齐次的, 并且 $\partial^M\phi(x^0, y^0)(0)$ 是闭锥.

例 8.4.1 设集合 $S \subset R^n$ 非空, 集合

$$\phi(x) = \begin{cases} 0, & x \in S; \\ \emptyset, & x \notin S. \end{cases}$$

考虑集值映射 $\phi: R^n \rightarrow 2^{R^n}$, $x \mapsto \phi(x)$, 由定义 8.4.2 则有

$$\partial^M\phi(x^0)(y^*) = N_S^M(x^0) \quad \forall x^0 \in S, y^* \in R^n.$$

定义 8.4.3 设集合 $S \subset R^n$ 非空, 点 $x^0 \in \text{cl}S$, $\epsilon \geq 0$, 则称集合

$$N_S^F(x^0, \epsilon) = \left\{ x^* \in R^n \mid \lim_{\substack{x' \rightarrow x^0 \\ x' \in S}} \sup \frac{\langle x^*, x' - x^0 \rangle}{\|x' - x^0\|} \leq \epsilon \right\} \quad (8.4.2)$$

是集合 S 在点 x^0 处的 F -法锥 (Frechet 的 ϵ -法锥). 若 $x^0 \notin \text{cl}S$, 令 $N_S^F(x^0, \epsilon) = \emptyset$, 当 $\epsilon = 0$ 时, 记 $N_S^F(x^0) = N_S^F(x^0, 0)$.

注 8.4.2 有下面关系:

$$N_S^F(x^0) = \limsup_{x \rightarrow x^0} N_S^F(x) = \limsup_{x \rightarrow x^0, \epsilon \rightarrow 0^+} N_S^F(x, \epsilon). \quad (8.4.3)$$

定义 8.4.4 设点 $x^0 \in R^n$, $y^0 \in \psi(x^0) \subset R^m$.

(1) 称

$$\begin{aligned} \partial^{MC}\psi(x^0, y^0)(y^*) \\ = \{x^* \in R^n \mid (x^*, -y^*) \in N_{\text{graph}\psi}^C(x^0, y^0)\} \end{aligned} \quad (8.4.4)$$

是集值映射 ψ 在点 (x^0, y^0) 处的 MC -次微分.

(2) 称

$$\begin{aligned} \partial^{MF}\psi(x^0, y^0)(y^*) \\ = \{x^* \in R^n \mid (x^*, -y^*) \in N_{\text{graph}\psi}^F(x^0, y^0)\} \end{aligned} \quad (8.4.5)$$

是集值映射 ψ 在点 (x^0, y^0) 处的 MF -次微分.

注 8.4.3 显然, 这里由 C -切锥和 F -法锥引进的 MC -次微分和 MF -次微分与定义 8.4.2 中给出的 M -次微分是不一样的.

注 8.4.4 我们知道, 当 S 是凸集时, 有 $N_S^C(x^0) = N_S^M(x^0)$. 若 $N_S^C(x^0) = N_S^M(x^0)$ 或 $N_S^M(x^0) = N_S^F(x^0)$ 时, 则称集合 S 在点 x^0 处是正则的. 显然, 若 $\text{graph}\psi$ 在点 (x^0, y^0) 处是正则的, 则 (8.4.1) 与 (8.4.4) 或 (8.4.1) 与 (8.4.5) 是一致的, 这时称 ψ 在点 (x^0, y^0) 处是可微正则的.

定理 8.4.1 设 $x \in R^n$, $y \in \psi(x)$. 若 ψ 在点 x 附近是单值的, 在点 x 处是可微的, $J\psi(x) \in R^{m \times n}$, 即

$$\lim_{x', x'' \rightarrow x} \frac{(\psi(x') - \psi(x'') - J\psi(x)(x' - x''))}{\|x' - x''\|} = 0.$$

则 ψ 在 (x, y) 处是 M -次可微的和正则的, 并且有

$$\partial^M\psi(x)(y^*) = \{(J\psi(x))^* y^*\} \quad \forall y^* \in R^m.$$

证明 由定理 8.3.3 和定义 8.4.2 不难推得. \square

定理 8.4.2 设 $\text{graph}\psi$ 是凸的, 则 ψ 在 $(x, y) \in \text{cl}(\text{graph}\psi)$ 处是可微和正则的, 并且对任意的 $y^* \in R^m$ 有

$$\begin{aligned} \partial^M \psi(x, y)(y^*) &= \{x^* \in R^n \mid \langle x^*, x \rangle - \langle y^*, y \rangle \\ &= \sup[\langle x^*, x' \rangle - \langle y^*, y' \rangle \quad \forall (x', y') \in \text{graph}\psi]\}. \end{aligned}$$

证明 由定理 8.3.4 和定理 6.2.11(当 $p=1$) 可得证. \square

定义 8.4.5 设点 $x^0 \in R^n$, $y^0 \in \psi(x^0)$, $B \subset R^m$ 是开单位球. 若存在点 x^0 的邻域 $U(x^0)$ 和点 y^0 的邻域 $U(y^0)$ 以及常数 $\gamma \geq 0$, 使得

$$\psi(x') \cap U(y^0) \subset \psi(x) + \gamma \|x^* - x\| B \quad \forall x, x' \in U(x^0),$$

则称集值映射 ψ 在点 $(x^0, y^0) \in \text{graph}\psi$ 处是伪 Lipschitz 的.

定理 8.4.3 设 $\text{graph}\psi$ 是闭的, $(x, y) \in \text{graph}\psi$, 则 ψ 在点 (x, y) 处是伪 Lipschitz 的当且仅当

$$\partial^M \psi(x, y)(0) = \{0\}.$$

证明 由定义 8.4.1 和定义 8.4.5 可得证. \square

设 ψ^{-1} 是 ψ 的逆映射: $\psi^{-1}(y) = \{x \in R^n \mid (x, y) \in \text{graph}\psi\}$, 则有

$$y^* \in \partial^M \psi^{-1}(y, x)(x^*) \iff -x^* \in \partial^M \psi(x, y)(-y^*).$$

定义 8.4.6 设 $f: R^n \rightarrow R \cup \{\pm \infty\}$ 是广义实值函数, 点 $x^0 \in \text{dom}f$.

(1) 称集合

$$\begin{aligned} \partial^M f(x^0) &= \partial^M \text{epi}f(x^0, f(x^0))(1) \\ &= \{x^* \in R^n \mid (x^*, -1) \in N_{\text{epi}f}^M(x^0, f(x^0))\} \end{aligned}$$

是 f 在点 x^0 处的 M -次微分. 当 $x^0 \notin \text{dom}f$, 则令 $\partial^M f(x^0) = \emptyset$.

(2) 称集合

$$\begin{aligned}\partial^{\infty} f(x^0) &= \partial^M \text{epi} f(x^0, f(x^0)) \setminus \{0\} \\ &= \{x^* \in R^n \mid (x^*, 0) \in N_{\text{epi} f}^M(x^0, f(x^0))\}\end{aligned}$$

是 f 在点 x^0 处的 M -奇异次微分.

注 8.4.5 注意, 这里在定义 8.4.6 中使用了实值函数的上图象. 在定义 8.4.2 中的集值映射退化为函数时, 这两个定义是不一致的. 当 f 在点 x^0 处是严格可微的, 则有 $\partial^M f(x^0) = \{\nabla f(x^0)\}$.

注 8.4.6 设 f 在点 x 处是 Lipschitz 的, $\partial^0 f(x)$ 是 f 在 x 处的 G -次微分, 则有

$$\partial^0 f(x) = \text{cl}(\text{co}[\partial^M f(x) + \partial^{\infty} f(x)])$$

注 8.4.7 设凸集 $S \subset R^n$, 对于 S 上的指示函数:

$$\delta_S(x) = \begin{cases} 0 & x \in S; \\ +\infty, & x \in R^n \setminus S, \end{cases}$$

则有

$$\partial^M \delta_S(x) = \partial^{\infty} \delta_S(x) = \partial_S(x) = N_S^M(x) \quad \forall x \in S.$$

定理 8.4.4 设 $x \in R^n$. 若向量函数 $f: R^n \rightarrow R^m$ 在点 x 处是 Lipschitz 的, 则

$$\partial^M f(x)(y^*) = \partial^M \langle y^*, f \rangle(x) \quad \forall y^* \in R^m.$$

证明 由定义 8.4.2 可得证. \square

为了讨论由定义 8.4.2 引进的 M -次微分的运算性质, 以下给出集合的局部极点的概念.

定义 8.4.7 设 S_1, S_2 是闭集, 点 $x^0 \in S_1 \cap S_2$. 若存在点 x^0 的邻域 $U(x^0)$ 及序列 $\{a_k^1\}$ 和 $\{a_k^2\}$, $a_k^1 \rightarrow 0$, $a_k^2 \rightarrow 0 (k \rightarrow \infty)$, 使得

$$(S_1 - a_k^1) \cap (S_2 - a_k^2) \cap U(x^0) = \emptyset \quad k = 1, 2, \dots,$$

则称 x^0 是集系 $\{S_1, S_2\}$ 的局部极点. 若 $\{S_1, S_2\}$ 至少有一个局部极点, 则称集合 S_1 和 S_2 形成一个局部极点系.

例如, 设 x^0 是 S 的边界点, 则 x^0 是 $\{x^0, S\}$ 的局部极点. 又如, 设函数 $f: R^n \rightarrow R$ 在点 x^0 处下半连续, $S \subset R^n$ 是闭凸集, x^0 是极小化问题

$$\min_{x \in S} f(x)$$

的局部最优解. 取 $S_1 = \text{epi} f$ 和 $S_2 = S \times \{f(x)\}$, 则点 $(x^0, f(x^0))$ 是 $\{S_1, S_2\}$ 的局部极点.

依据上面所述, 局部极点可以看成是局部最优解的推广.

下面给出局部极点的必要条件.

定理 8.4.5 设 S_1, S_2 是闭集. 若 $\bar{x} \in S_1 \cap S_2$ 是 $\{S_1, S_2\}$ 的局部极点, 则存在 $x^* \in R^n$, 使得

$$0 \neq x^* \in N_{S_1}(\bar{x}) \cap (-N_{S_2}(\bar{x})). \quad (8.4.6)$$

证明 考虑序列 $\{a_k^1\}$ 和 $\{a_k^2\}$, $a_k^1 \rightarrow 0$, $a_k^2 \rightarrow 0$ ($k \rightarrow \infty$), 设 $U(\bar{x})$ 是 \bar{x} 的邻域, 则有

$$(S_1 - a_k^1) \cap (S_2 - a_k^2) \cap U(\bar{x}) = \emptyset, \quad k = 1, 2, \dots$$

用 $d_S(x + a)$ 表示点 $x + a$ 到集合 S 的距离, 设

$$\rho_k(x) = [d_{S_1}^2(x + a_k^1) + d_{S_2}^2(x + a_k^2)]^{1/2} + \|x - \bar{x}\|^2,$$

考虑一系列极小化问题:

$$\min_{x \in R^n} \rho_k(x), \quad k = 1, 2, \dots. \quad (8.4.7)$$

因为对任何 k , $\rho_k(x)$ 关于 x 是连续的, 并且水平集 $\{x | \rho_k(x) \geq 0\}$ 有界, 所以对于每一个 k , 存在 (8.4.7) 的最优解 x^k . 设

$$\alpha_k = [d_{S_1}^2(x^k + a_k^1) + d_{S_2}^2(x^k + a_k^2)]^{1/2},$$

则有

$$\begin{aligned}\rho_k(x^k) &= \alpha_k + \|x^k - \bar{x}\|^2 \leq \rho_k(\bar{x}) \\ &\leq [\|a_k^1\|^2 + \|a_k^2\|^2]^{1/2}.\end{aligned}$$

由于已知当 $k \rightarrow \infty$ 时, 有 $a_k^1 \rightarrow 0$ 和 $a_k^2 \rightarrow 0$, 故得

$$\alpha_k \rightarrow 0^+ \quad \text{和} \quad x^k \rightarrow \bar{x}, \quad k \rightarrow \infty. \quad (8.4.8)$$

设 $\omega_k^1 \in d_{S_1}(x^k + a_k^1)$ 和 $\omega_k^2 \in d_{S_2}(x^k + a_k^2)$, 考虑极小化问题

$$\begin{aligned}\min_{x \in R^n} f_k(x) &= [\|x + a_k^1 - \omega_k^1\|^2 + \|x + a_k^2 - \omega_k^2\|^2]^{1/2} \\ &\quad + \|x - \bar{x}\|^2 \quad k = 1, 2, \dots.\end{aligned}$$

显然, 上式有最优解 $x^k (k = 1, 2, \dots)$. 由 $f_k(x) (k = 1, 2, \dots)$ 在 x 处是连续可微的, 则

$$\nabla f_k(x^k) = x_1^{k*} + x_2^{k*} + 2(x^k - \bar{x}) = 0, \quad (8.4.9)$$

其中 $x_i^{k*} = (x^k + a_k^i - \omega_k^i)/\alpha_k (i = 1, 2)$, 并且 $\|x_1^{k*}\|^2 + \|x_2^{k*}\|^2 = 1$. 因此, 存在 x_1^* 和 x_2^* , 使得

$$\|x_1^{k*}\|^2 + \|x_2^{k*}\|^2 = 1, \quad x_1^{k*} \rightarrow x_1^* \quad \text{和} \quad x_2^{k*} \rightarrow x_2^*, \quad k \rightarrow \infty.$$

对(8.4.9)取极限, 根据定义 8.4.1 即得(8.4.6). \square

注 8.4.8 若 S_1 和 S_2 是凸集, 则(8.4.6)等价于凸集分离定理, 即存在 $x^* \neq 0$, 使得

$$\langle x^*, x^1 \rangle \leq \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, x^2 \in S_2.$$

事实上, 由凸集分离性可以推出局部极点性. 在两个集合具有凸性时, 凸集分离性与局部极点性是等价的.

定理 8.4.6 设集合 $\phi_1(x), \phi_2(x) \subset R^m (x \in R^n)$, 集值映射 $\phi_1: R^n \rightarrow 2^{R^m}$ 和 $\phi_2: R^n \rightarrow 2^{R^m}$ 是闭的, $y^0 \in \phi_1(x^0) + \phi_2(x^0)$. 若集值映射 $\Theta: R^{n+m} \rightarrow 2^{R^{2m}}$,

$$\begin{aligned}\Theta(x, y) &= \{(y_1, y_2) \in R^{2m} \mid y_1 \in \phi_1(x), \\ &\quad y_2 \in \phi_2(x), y_1 + y_2 = y\} \quad (8.4.10)\end{aligned}$$

在 (x^0, y^0) 处局部有界, 并且下式成立:

$$\begin{aligned} \partial^M \phi_1(x^0, y_1)(\theta) \cap (-\partial^M \phi_2(x^0, y_2)(\theta)) &= \{\theta\} \\ \forall (y_1, y_2) &\in \Theta(x^0, y^0), \end{aligned} \quad (8.4.11)$$

则对任意的 $y^* \in R^m$ 有

$$\begin{aligned} &\partial^M(\phi_1 + \phi_2)(x^0, y^0)(y^*) \\ &\subset \bigcup_{(y_1, y_2) \in \Theta(x, y)} [\partial^M \phi_1(x^0, y_1)(y^*) + \partial^M \phi_2(x^0, y_2)(y^*)]. \end{aligned} \quad (8.4.12)$$

证明 设 $x^* \in \partial^M(\phi_1 + \phi_2)(x^0, y^0)(y^*)$, 由 ϕ_1 和 ϕ_2 是闭的可知, $\phi_1 + \phi_2$ 的图象也是闭的. 再由定义 8.4.2 和 F -法锥的关系得知, 对任何一个正数列 $\{\epsilon_k\}$, $\epsilon_k \rightarrow 0^+$ ($k \rightarrow \infty$), 都可以找到序列 $\{x^k\}$ 和 $\{y^k\}$, $\{x_k^*\}$ 和 $\{y_k^*\}$, $x^k \rightarrow x^0$, $y^k \rightarrow y^0$, $x_k^* \rightarrow x^*$, $y_k^* \rightarrow y^*$ ($k \rightarrow \infty$), 以及正数列 $\{\eta_k\}$, $\eta_k \rightarrow 0^+$ ($k \rightarrow \infty$), 使得 $y^k \in \phi_1(x^k) + \phi_2(x^k)$ 和对任意的 $(x, y) \in \text{graph}(\phi_1 + \phi_2)$, 当 $\|x - x^k\| \leq \eta_k$, $\|y - y^k\| \leq 2\eta_k$ 时有

$$\begin{aligned} &\langle x^k, x - x^k \rangle - \langle y_k^*, y - y^k \rangle \\ &\leq \epsilon_k (\|x - x^k\| + \|y - y^k\|), \end{aligned} \quad (8.4.13)$$

由于 $\Theta(x, y)$ 在 (x^0, y^0) 处是局部有界的, 故存在序列 $\{(y_1^k, y_2^k)\} \subset \Theta(x^k, y^k)$ 收敛于 $(y_1^0, y_2^0) \in \Theta(x^0, y^0)$. 对每一个 $k = 1, 2, \dots$, 设

$$\begin{aligned} g_k^1(x, y_1, y_2) &= \langle y_k^*, y_1 - y_1^k \rangle + \epsilon_k \|y_1 - y_1^k\| \\ &\quad + \delta((x_1, y_1), \text{graph} \phi_1), \end{aligned} \quad (8.4.14)$$

$$\begin{aligned} g_k^2(x, y_1, y_2) &= \langle x_k^*, x - x^k \rangle - \epsilon_k \|x - x^k\| \\ &\quad - \langle y_k^*, y_2 - y_2^k \rangle - \epsilon_k \|y_2 - y_2^k\| \\ &\quad - \delta((x, y_2), \text{graph} \phi_2), \end{aligned} \quad (8.4.15)$$

由(8.4.13)和(8.4.14)作相应于 $R^n \times R^m \times R^m \times R$ 的闭子集:

$$S_k^1 = \text{epig}_k^1, S_k^2 = \text{epig}_k^2, \quad (8.4.16)$$

显然有

$$(x^k, y_k^1, y_k^2, 0) \in S_k^1 \cap S_k^2, \quad k = 1, 2, \dots$$

由(8.4.13)得

$$S_k^1 \cap [S_k^2 - (0, 0, 0, v)] \cap [B((x^k, y_k^1, y_k^2), \eta_k) \times R] = \emptyset \\ \forall v > 0,$$

故由定义 8.4.4 可知 $(x^k, y_k^1, y_k^2, 0)$ 是 (S_k^1, S_k^2) 的局部极点. 由此, 根据定理 8.4.5 得知, 存在序列 $\{z_k^*, y_k^{1*}, y_k^{2*}, \lambda_k\} \in R^{n+2m+1}$, 使得

$$\|(z_k^*, y_k^{1*}, y_k^{2*}, \lambda_k)\| = 1, \quad k = 1, 2, \dots \quad (8.4.17)$$

和

$$(z_k^*, y_k^{1*}, y_k^{2*}, \lambda_k) \in N_{S_k^1}^M(x^k, y_k^1, y_k^2, 0) \\ \cap [-N_{S_k^2}^M(x_k, y_k^1, y_k^2, 0)].$$

由(8.4.14)和(8.4.16)得

$$(z_k^*, y_k^{1*}, y_k^{2*}, \lambda_k) \in N_{\text{epig}_k^1}^M(x^k, y_k^1, y_k^2, g_k^1(x^k, y_k^1, y_k^2)). \quad (8.4.18)$$

根据定义 8.4.1 和定义 8.4.3, 上式等价于

$$(z_k^*, y_k^{1*}, y_k^{2*}) \in \begin{cases} |\lambda_k| \partial^M g_k^1(x^k, y_k^1, y_k^2), & \lambda_k \neq 0; \\ \partial^\infty g_k^1(x^k, y_k^1, y_k^2), & \lambda_k = 0. \end{cases}$$

由此, 从上式可以推得 $y_k^{2*} = 0$,

$$(z_k^*, y_k^{1*})$$

$$\in \begin{cases} (0, |\lambda_k|(y_k^* + \epsilon_k B)) + N_{\text{graph } \phi_1}^M(x^k, y_k^1), & \lambda_k \neq 0; \\ N_{\text{graph } \phi_1}^M(x^k, y_k^1), & \lambda_k = 0. \end{cases} \quad (8.4.19)$$

同样地,我们有

$$\begin{aligned} & (-z^*, -y_1^{1*}, -y_1^{2*}, -\lambda_k) \\ & \in N_{\text{prox}_k^2}^M((x^k, y_1^1, y_1^2), g_k^2(x^k, y_1^1, y_1^2)). \end{aligned}$$

从上式可以推得 $y_i^{1*} = 0$ 和

$$\begin{aligned} (z_k^*, y_k^{2*}) = & \begin{cases} (|\lambda_k|(x_k^* - \varepsilon_k B, -y_k^* - \varepsilon_k B)) - N_{\text{graph}\phi_2}^M(x^k, y_k^2), & \lambda_k \neq 0; \\ -N_{\text{graph}\phi_2}^M(x^k, y_k^2), & \lambda_k = 0. \end{cases} \end{aligned} \quad (8.4.20)$$

由(8.4.17), $y_i^{1*} = 0$ 和 $y_i^{2*} = 0$, 可选出序列 $\{z_k^*, \lambda_k\}$, 使得当 $k \rightarrow \infty$ 时, 它收敛于 $(z^*, \lambda) \in R^n \times R$:

$$\|(z^*, \lambda)\| = 1, \lambda \leq 0. \quad (8.4.21)$$

我们证明 $\lambda \neq 0$. 事实上, 否则, 若 $\lambda = 0$, 则有 $z^* \neq 0$, 由(8.4.21), (8.4.19)和(8.4.20), 根据定义 8.4.2 我们有

$$z^* \in \partial^M \phi_1(x^0, y_1)(0) \cap (-\partial^M \phi_2(x^0, y_2)(0))$$

$$\forall (y_1^0, y_2^0) \in \Theta(x^0, y^0),$$

这导致与(8.4.11)矛盾. 因此, 由(8.4.21)有 $\lambda < 0$. 对(8.4.19)取极限, 得

$$z^* |\lambda|^{-1} \in \partial^M \phi_1(x^0, y_1^0)(y^*),$$

对(8.4.20)取极限, 则得

$$(x^*, -z^*) |\lambda|^{-1} \in \partial^M \phi_2(x^0, y_2^0)(y^*).$$

由此证明了(8.4.12). \square

推论 8.4.7 若对每一个 $(y_1, y_2) \in \Theta(x^0, y^0)$, 集值映射 ϕ_1 在 (x^0, y_1) 处或 ϕ_2 在 (x^0, y_2) 处是伪 Lipschitz 的, Θ 在 (x^0, y^0) 处是局部有界的, 则(8.4.12)成立.

证明 由定理 8.4.3 和定理 8.4.6 可以推得结论. \square

推论 8.4.8 设向量函数 $f_1: R^n \rightarrow R^m$ 在 $x^0 \in R^n$ 处是可微的, 集值映射 $\phi_2: R^n \rightarrow 2^{R^m}$ 是闭的, 对任意的 $y^0 \in f_1(x^0) + \phi_2(x^0)$ 和 $y^* \in R^m$, 则有

$$\begin{aligned} \partial^M(f_1 + \phi_2)(x^0, y^0)(y^*) \\ = (Jf_1(x^0) \mid y^* + \partial^M\phi_2(x^0, y^0 - f_1(x^0))(y^*)). \end{aligned}$$

并且, $f_1 + \phi_2$ 在 (x^0, y^0) 处是正则和可微的当且仅当 ϕ_2 在 $(x^0, y^0 - f_1(x^0))$ 处是正则和可微的.

证明 由定理 8.4.1 和定理 8.4.6 可得证. \square

推论 8.4.9 设集值映射 $\phi_1: R^n \rightarrow 2^{R^m}$ 是闭的, 又 $S \in R^n$ 是闭集,

$$\phi_2(x) = \begin{cases} 0 \in R^m, & x \in S; \\ \emptyset, & x \notin S. \end{cases} \quad (8.4.22)$$

给定 $x^0 \in S, y^0 \in \phi_1(x^0)$, 设

$$\partial^M\phi_1(x^0, y^0)(\theta) \cap (-N_S^M(x^0)) = \{\theta\}, \quad (8.4.23)$$

则有

$$\begin{aligned} \partial^M(\phi_1 + \phi_2)(x^0, y^0)(y^*) \subset \partial^M\phi_1(x^0, y^0)(y^*) + N_S^M(x^0) \\ \forall y^* \in R^m. \end{aligned} \quad (8.4.24)$$

再若 ϕ_1 在 x^0 处是可微的或在 (x^0, y^0) 处是正则和可微的, 并且 S 在 x^0 处是正则的, 则 (8.4.24) 成为等式.

证明 设 $\Theta(x^0, y^0) = \{(y^0, \theta)\}$. 由 (8.4.22) 和 (8.4.23), 根据定理 8.4.6 可推知 (8.4.24) 成立.

若 ϕ_1 在 x^0 处是严格可微的, 据推论 8.4.8 可知 (8.4.24) 为等式. 若 ϕ_1 在 (x^0, y^0) 处是正则和可微的, 以及 S 在 x^0 处是正则的, 则有

$$\partial^M\phi_1(x^0, y^0)(y^*) + N_S^M(x^0)$$

$$\begin{aligned}
 &= \partial^{MF} \phi_1(x^0, y^0)(y^*) + N_S^F(x^0) \\
 &\subset \partial^{MF}(\phi_1 + \phi_2)(x^0, y^0)(y^*) \\
 &\subset \partial^M(\phi_1 + \phi_2)(x^0, y^0)(y^*) \\
 &\quad \forall y^* \in R^m.
 \end{aligned}$$

因此, (8.4.24) 成为等式. \square

推论 8.4.10 设函数 $f_i: R^n \rightarrow R$ ($i = 1, 2$) 是下半连续的. 若对 $x \in R^n$, 有

$$\partial^\infty f_1(x) \cap (-\partial^\infty f_2(x)) = \{\theta\}, \quad (8.4.25)$$

则

$$\partial^M(f_1 + f_2)(x) \subset \partial^M f_1(x) + \partial^M f_2(x), \quad (8.4.26)$$

$$\partial^\infty(f_1 + f_2)(x) \subset \partial^\infty f_1(x) + \partial^\infty f_2(x). \quad (8.4.27)$$

再若 f_1 或 f_2 在 x 处是可微的, 则 (8.4.26) 和 (8.4.27) 成为等式.

证明 由定理 8.4.4 和定理 8.4.6 可得证. \square

推论 8.4.11 设 S_1, S_2 是闭集. 若 $x \in S_1 \cap S_2$, $N_{S_1}^M(x) \cap (-N_{S_2}^M(x)) = \{\theta\}$, 则

$$N_{S_1 \cap S_2}^M(x) \subset N_{S_1}^M(x) + N_{S_2}^M(x).$$

证明 由定义 8.4.1, 例 8.4.1 和定理 8.4.6 可以推得. \square

下面讨论复合集值映射的 M -次微分问题.

设集合 $\theta(x) \subset R^m$ ($x \in R^n$), $\phi(y) \subset R^q$ ($y \in R^m$, $q > 1$), 对于集值映射 $\theta: R^n \rightarrow 2^{R^m}$ 和 $\phi: R^m \rightarrow 2^{R^q}$, 考虑复合集值映射

$$(\phi \circ \theta): R^n \rightarrow 2^{R^q},$$

$$x \mapsto (\phi \circ \theta)(x) = \phi(\theta(x)) = \bigcup_{y \in \theta(x)} \phi(y).$$

定理 8.4.12 设 $x^0 \in R^n$, $\text{graph} \theta$ 和 $\text{graph} \phi$ 是闭的, $z^0 \in (\phi \circ \theta)(x^0)$. 考虑集值映射 $E: R^n \times R^q \rightarrow 2^{R^m}$, $(x, z) \mapsto E(x, z)$,

$$\Xi(x, z) = \theta(x) \cap \psi^{-1}(z) = \{y \in \theta(x) \mid z \in \theta(y)\}. \quad (8.4.28)$$

若 Ξ 在点 (x^0, z^0) 处是局部有界的, 且条件

$$\begin{aligned} \partial^M \psi(y, z^0)(\theta) \cap \text{Ker} \partial^M \theta(x^0, y) &= \{\theta\} \\ \forall y \in \theta(x^0) \cap \psi^{-1}(z^0) \end{aligned} \quad (8.4.29)$$

成立, 则

$$\begin{aligned} \partial^M(\psi \circ \theta)(x^0, z^0) \\ \subset \bigcap_{y \in \theta(x^0) \cap \psi^{-1}(z^0)} [\partial^M \theta(x^0, y) \circ \partial^M \psi(y, z^0)]. \end{aligned} \quad (8.4.30)$$

证明 设 $x^* \in \partial^M(\psi \circ \theta)(x^0, z^0)(z^*)$, 我们证明存在 $y \in \theta(x^0) \cap \psi^{-1}(z^0)$ 和 $y^* \in \partial^M \psi(y, z^0)(z^*)$, 使得 $x^* \in \partial^M \theta(x^0, y)(y^*)$. 为此, 作集值映射 $g: R^n \times R^m \rightarrow 2^{R^n}$, $(x, y) \mapsto g(x, y)$,

$$g(x, y) = \psi(y) + \Delta^g((x, y), \text{graph} \theta), \quad (8.4.31)$$

其中

$$\Delta^g((x, y), \text{graph} \theta) = \begin{cases} 0, & (x, y) \in \text{graph} \theta; \\ \emptyset, & (x, y) \notin \text{graph} \theta. \end{cases}$$

对任意的 $z^* \in R^q$ 我们证明

$$\begin{aligned} \partial^M(\psi \circ \theta)(x^0, z^0)(z^*) \\ \subset \left\{ x^* \mid (x^*, \theta) \in \bigcup_{y \in \theta(x^0) \cap \psi^{-1}(z^0)} \partial^M g(x^0, y, z^0)(z^*) \right\}. \end{aligned} \quad (8.4.32)$$

设 $x^* \in \partial^M(\psi \circ \theta)(x^0, z^0)(z^*)$, 由于 Ξ 在点 (x^0, z^0) 处是局部有界的, 则 $\text{graph}(\psi \circ \theta)$ 是闭的. 根据定义 8.4.2 和 (8.4.3) 式可知, 存在序列 $\{x^k\}$ 和 $\{z^k\}$, 以及 $\{x_k^*\}$ 和 $\{z_k^*\}$, $x^k \rightarrow x^0$, $z^k \rightarrow z^0$, 以及 $x_k^* \rightarrow x^*$, $z_k^* \rightarrow z^*$ ($k \rightarrow \infty$), 使得

$$z^k \in (\psi \circ \theta)(x^k) \text{ 和 } (x_k^*, -z_k^*) \in N_{\text{graph}(\psi \circ \theta)}^M(x^k, z^k),$$

$$k = 1, 2, \dots,$$

取任一序列 $\{y^k\}$, $y^k \in \theta(x^k) \cup \psi^{-1}(z^k)$, 又由于 Ξ 在点 (x^0, z^0) 处是局部有界的, 故不妨设 $y_k^0 \rightarrow y^0 \in \Xi(x^0, z^0)$. 由 (8.4.2) 令 $\varepsilon = 0$, 我们有

$$\begin{aligned} \gamma &= \limsup_{\substack{(x, y, z) \rightarrow (x^k, y^k, z^k) \\ y \in \theta(x), z \in \psi(y)}} \frac{\langle x_k^*, x - x^k \rangle - \langle z_k^*, z - z^k \rangle}{\|(x, y, z) - (x^k, y^k, z^k)\|} \\ &\leq \limsup_{\substack{(x, z) \rightarrow (x^k, z^k) \\ z \in (\psi, \theta)(x)}} \frac{\langle x_k^*, x - x^k \rangle - \langle z_k^*, z - z^k \rangle}{\|(x, z) - (x^k, z^k)\|} \leq 0. \end{aligned}$$

由此, 得知 $(x_k^*, 0 - z_k^*) \in N_{\text{graph } \theta}^M(x^k, y^k, z^k)$, 再由 (8.4.3) 得到 (8.4.32). 现在由 (8.4.31) 利用推论 8.4.9, 即可推得 (8.4.30) 成立. \square

推论 8.4.13 在定理 8.4.12 中, 条件 (8.4.31) 可以换成: 对任意的 $y \in \theta(x^0) \cup \psi^{-1}(z^0)$, Ξ 在 (y, z^0) 处是伪 Lipschitz 的.

证明 由定理 8.4.3 和定理 8.4.12 可得证. \square

推论 8.4.14 设向量函数 $f: R^m \rightarrow R^q$ 在点 $y \in \theta(x^0) \cap f^{-1}(z^0)$ 处是可微的, $\text{graph } \theta$ 是闭的. 若由 (8.4.28) 定义的 Ξ 在点 (x^0, z^0) 处是局部有界的, 则

$$\begin{aligned} &\partial^M(f \circ \theta)(x^0, z)(z^*) \\ &\subset \bigcup_{y \in \theta(x^0) \cap f^{-1}(z^0)} [\partial^M \theta(x^0, y) \circ (Jf(y))^*]. \end{aligned}$$

证明 由定理 8.4.1 和定理 8.4.12 可以得证. \square

推论 8.4.15 设向量函数 $f: R^m \rightarrow R^m$ 在点 x^0 处是可微的, $\text{graph } \psi$ 是闭的. 若

$$\partial^M \psi(f(x^0), z^0)(\theta) \cap \text{Ker}(Jf(x^0))^* = \{\theta\},$$

则对任意的 $z^* \in R^q$ 有

$$\begin{aligned} & \partial^M(\psi \circ f)(x^0, z^0)(z^*) \\ & \subset (Jf(x^0))^* \partial^M \psi(f(x^0), z^0)(z^*). \end{aligned}$$

证明 由定理 8.4.1 和定理 8.4.12 可得证. \square

定理 8.4.16 设向量函数 $\theta: R^n \rightarrow R^m$ 在 x^0 处是 Lipschitz 连续的, 集值映射 $\psi: R^m \rightarrow 2^{R^q}$ 是闭的, 则由 (8.4.31) 定义的 g 使 (8.4.32) 成为等式. 再若条件 (8.4.29) 成立, $y^0 = \theta(x^0)$, 则

$$\begin{aligned} & \partial^M(\psi \circ \theta)(x^0, z^0)(z^*) \\ & \subset \bigcup_{y^* \in \partial^M \psi(\theta(x^0), z^0)(z^*)} \partial^M \psi(y^*, \theta)(x^0). \end{aligned} \quad (8.4.33)$$

(1) 若 ψ 在 $y^0 = \theta(x^0)$ 处是可微的, 则

$$\begin{aligned} & \partial^M(\psi \circ \theta)(x^0, z^0)(z^*) \\ & = \partial^M \langle (J\psi(y^0))^* z^*, \theta \rangle(x^0) \quad \forall z^* \in R^q. \end{aligned} \quad (8.4.34)$$

(2) 若 θ 在 x^0 处是可微的, ψ 在 $(\theta(x^0), z^0)$ 处是可微和正则的, 则有

$$\begin{aligned} & \partial^M(\psi \circ \theta)(x^0, z^0)(z^*) \\ & = (J\theta(x^0))^* \partial^M \psi(\theta(x^0), z^0)(z^*) \quad \forall z^* \in R^q, \end{aligned} \quad (8.4.35)$$

并且 $\theta \circ \psi$ 在 (x^0, z^0) 处是可微和正则的.

证明 先证明 (8.4.32) 的反包含情形: 设 (x^*, θ) 属于 (8.4.32) 的右端, 即

$$(x^*, \theta, -z^*) \in N_{\text{graph}g}^M(x^0, \theta(x^0), z^0).$$

由 (8.4.3) 可知, 存在序列 $\{x^k\}$ 、 $\{z^k\}$ 、 $\{x_k^*\}$ 、 $\{y_k^*\}$ 和 $\{z_k^*\}$, $x^k \rightarrow x^0$, $z^k \rightarrow z^0$, $x_k^* \rightarrow x^*$, $y_k^* \rightarrow \theta$ 和 $z_k^* \rightarrow z^*$ ($k \rightarrow \infty$), 使得

$$\begin{aligned} & \limsup_{\substack{(x, z) \rightarrow (x^k, z^k) \\ z \in (\psi \circ \theta)(x)}} \frac{\langle (x_k^*, y_k^*, -z_k^*), (x, \theta(x), z) - (x^k, \theta(x^k), z^k) \rangle}{\|(x, \theta(x), z) - (x^k, \theta(x^k), z^k)\|} \leq 0, \\ & k = 1, 2, \dots, \end{aligned} \quad (8.4.36)$$

设 γ 为 θ 在点 x^0 处的 Lipschitz 常数, 则有

$$\begin{aligned} & \| (x, \theta(x), z) - (x^k, \theta(x^k), z^k) \| \\ & \leq (\gamma + 1) (\|x - x^k\| + \|z - z^k\|). \end{aligned}$$

因此有

$$\begin{aligned} & \frac{\langle (x_k^*, -z_k^*), (x, z) - (x^k, z^k) \rangle}{\|(x, z) - (x^k, z^k)\|} \\ & \leq (\gamma + 1) \frac{\langle (x_k^*, -z_k^*), (x, z) - (x^k, z^k) \rangle}{\|(x, \theta(x), z) - (x^k, \theta(x^k), z^k)\|} \\ & \leq (\gamma + 1) \frac{\langle (x_k^*, y_k^*, -z_k^*), (x, \theta(x), z) - (x^k, \theta(x^k), z^k) \rangle}{\|(x, \theta(x), z) - (x^k, \theta(x^k), z^k)\|} \\ & \quad + (\gamma + 1) \|y_k^*\|. \end{aligned}$$

设 $\varepsilon_k = (\gamma + 1) \|y_k^*\|$, 则(8.4.36)为

$$\limsup_{\substack{(x, z) \rightarrow (x^k, z^k) \\ x \in (\psi \circ \theta)(x)}} \frac{\langle x_k^*, x - x^k \rangle - \langle z_k^*, z - z^k \rangle}{\|(x, z) - (x^k, z^k)\|} \leq \varepsilon_k, \quad k = 1, 2, \dots.$$

上式等价于 $(x_k^*, -z_k^*) \in N_{\text{graph}(\psi \circ \theta)}^F((x_k, z_k), \varepsilon_k)$. 令 $\varepsilon_k \rightarrow 0^+$, 由(8.4.3)得 $x^* \in \partial^M(\psi \circ \theta)(x^0, z^0)(z^*)$, 再由定理 8.4.4 得到(8.4.33).

(1) 由推论 8.4.7 可得(8.4.34)成立.

(2) 由推论 8.4.8 可推得(8.4.35)成立. \square

推论 8.4.17 设 θ 和 ψ 都是单值映射(向量函数). 若它们分别在 x^0 和 $\theta(x^0)$ 处是 Lipschitz 连续的, 则对任意的 $z^* \in R^q$ 有

$$\partial^M(z^*, \psi \circ \theta)(x^0) \subset \bigcup_{y^* \in \partial^M(x^*, \psi)(\theta(x^0))} \partial^M(y^*, \theta)(x^0).$$

再若 ψ 在 $\theta(x^0)$ 处是严格可微的, 则上式成为等式.

证明 由定理 8.4.16 可推证. \square

推论 8.4.18 设 θ 和 ψ 与推论 8.4.17 中的相同, 则对任意的 $z^* \in R^q$ 有

$$\begin{aligned} (J(\psi \circ \theta)(x^0))^* z^* &\subset \text{co} \left\{ \bigcup_{y^* \in \partial^M(z^*, \psi)(\theta(x^0))} \partial^M \langle y^*, \theta(x^0) \rangle \right\} \\ &\subset \text{co} \left\{ \bigcup_{y^* \in (J\psi(\theta(x^0)))^* z^*} \partial^M \langle y^*, \theta(x^0) \rangle \right\}, \\ (J(\psi \circ \theta)(x^0))^* z^* &\subset \text{co} \{ (J\theta(x^0))^* \partial^M \langle z^*, \psi \rangle (\theta(x^0)) \} \\ &\subset \text{co} \{ (J\psi(\theta(x^0))J\theta(x^0))^* z^* \}, \end{aligned}$$

若 ψ 在 $\theta(x^0)$ 处是可微的, 则

$$(J(\psi \circ \theta)(x^0))^* z^* = (J\psi(\theta(x^0))J\theta(x^0))^* z^*.$$

证明 由定理 8.4.16 不难推证. \square

现在讨论乘法规则. 设集合 $\theta_1(x), \theta_2(x) \subset R^m (x \in R^n)$, $\theta_1: R^n \rightarrow 2^{R^m}$ 和 $\theta_2: R^n \rightarrow 2^{R^m}$ 是集值映射, 记

$$\langle \theta_1, \theta_2 \rangle(x) = \bigcup_{y_1 \in \theta_1(x), y_2 \in \theta_2(x)} \langle y_1, y_2 \rangle.$$

定理 8.4.19 设 $x \in R^n$, $\text{graph}\theta_1$ 和 $\text{graph}\theta_2$ 是闭的, $z \in \langle \theta_1, \theta_2 \rangle(x)$. 若集值映射 $E: R^n \times R \rightarrow 2^{R^m}$, $(x, z) \mapsto E(x, z)$,

$$E(x, z) = \{ (y^1, y^2) \in R^{2m} \mid y^1 \in \theta_1(x),$$

$$y^2 \in \theta_2(x), \langle y^1, y^2 \rangle = z \}$$

在 (x, z) 处是局部有界的, 则对任意的 $\alpha \in R$, 有

$$\begin{aligned} &\partial^M \langle \theta_1, \theta_2 \rangle(x, z)(\alpha) \\ &\subset \bigcup_{(y^1, y^2) \in E(x, z)} \partial^M(\theta_1, \theta_2)(x, y^1, y^2)(\alpha y^2, \alpha y^1). \end{aligned} \quad (8.4.37)$$

若 θ_1 和 θ_2 是单值映射, 并且在点 x 处是 Lipschitz 的, 则

$$\partial^M \langle \theta_1, \theta_2 \rangle(x)$$

$$= \partial^M \{ \langle \theta_2(x), \theta_1 \rangle + \langle \theta_1(x), \theta_2 \rangle \} (x). \quad (8.4.38)$$

证明 将集值映射 $\langle \theta_1, \theta_2 \rangle$ 看作下面两个集值映射的复合:

$$\theta(x) = \{ \theta_1(x), \theta_2(x) \} : R^n \rightarrow 2^{R^{2m}},$$

$$\psi(y^1, y^2) = \langle y^1, y^2 \rangle : R^{2m} \rightarrow 2^R,$$

即

$$\langle \theta_1, \theta_2 \rangle(x) = (\psi \circ \theta)(x).$$

根据推论 8.4.14 得知 (8.4.37) 成立. 再由定理 8.4.16 和定理 8.4.4 推得 (8.4.38) 成立. \square

推论 8.4.20 设 θ_1 或者 θ_2 是单值映射 (向量函数), 并且在点 $x \in R^n$ 处是 Lipschitz 的, 则

$$\partial^M \langle \theta_1, \theta_2 \rangle(x) \subset \partial^M \langle \theta_2(x), \theta_1 \rangle(x) + \partial^M \langle \theta_1(x), \theta_2 \rangle(x)$$

$$\subset \{ J\theta_1(x) \}^* \theta_2(x) + \{ J\theta_2(x) \}^* \theta_1(x).$$

再若 θ_1 或 θ_2 在 x 处是可微的, 则上式成为等式.

证明 由定理 8.4.1 和定理 8.4.19 可推得. \square

定理 8.4.21 题设与定理 8.4.19 相同. 若

$$\partial^M \theta_1(x, y^1)(\theta) \cap (-\partial^M \theta_2(x, y^2)(\theta)) = \{\theta\}$$

$$\forall (y^1, y^2) \in E(x, z),$$

则对任何实数 α 有

$$\partial^M \langle \theta_1, \theta_2 \rangle(x, z)(\alpha)$$

$$\subset \bigcup_{(y^1, y^2) \in E(x, z)} [\partial^M \theta_1(x, y^1)(\alpha y^2) + \partial^M \theta_2(x, y^2)(\alpha y^1)].$$

(8.4.39)

证明 作复合映射:

$$\langle \theta_1, \theta_2 \rangle(x) = (\psi \circ \theta)(x),$$

$$\theta(x) = (x, x); R^n \rightarrow 2^{R^{2n}},$$

$$\phi(x^1, x^2) = \langle \theta_1(x^1), \theta_2(x^2) \rangle; R^{2n} \rightarrow 2^R.$$

利用推论 8.4.14 可以推得(8.4.39). \square

定理 8.4.22 设集合 $\theta_1(x), \theta_2(x) \subset R (x \in R^n)$, $\theta_1, \theta_2: R^n \rightarrow 2^R$ 是集值映射, $\text{graph}\theta_1$ 和 $\text{graph}\theta_2$ 是闭凸的. 考虑集值映射:

$$\left\{ \frac{\theta_1}{\theta_2} \right\}(x) = \bigcup_{y_1 \in \theta_1(x), y_2 \in \theta_2(x)} \{y_1/y_2 \mid y_2 \neq 0\}.$$

设 $z \in \left\{ \frac{\theta_1}{\theta_2} \right\}(x)$. 若集值映射 $\Xi: R^n \times R \rightarrow 2^{R^2}$, $(x, z) \mapsto \Xi(x, z)$,

$$\Xi(x, z) = \{(y_1, y_2) \in R^2 \mid y_1 \in \theta_1(x),$$

$$y_2 \in \theta_2(x), y_1/y_2 = z\}$$

在 (x, z) 处是局部有界的, 则对任意的 $\alpha \in R$ 有

$$\begin{aligned} \partial^M \left\{ \frac{\theta_1}{\theta_2} \right\}(x, z)(\alpha) \subset & \bigcup_{(y_1, y_2) \in \Xi(x, z)} \left[\partial^M(\theta_1, \theta_2) \right. \\ & \left. \times (x, y_1, y_2)(\alpha y_1 - \alpha y_2) \right] / (y_2)^2. \end{aligned} \quad (8.4.40)$$

若 θ_1 和 θ_2 在 x 处是单值和 Lipschitz 的, 则对任意的 $\alpha \in R$ 有

$$\begin{aligned} \partial^M \left\{ \frac{\theta_1}{\theta_2} \right\}(x, z)(\alpha) \subset & \bigcup_{(y_1, y_2) \in \Xi(x, z)} \left[\partial^M \theta_1(x, y_1)(\alpha y_2) \right. \\ & \left. \times \partial^M \theta_2(x, y_2)(-\alpha y_1) \right] / (y_2)^2. \end{aligned} \quad (8.4.41)$$

证明 作复合集值映射: $\left\{ \frac{\theta_1}{\theta_2} \right\}(x) = (\psi \circ \theta)(x)$, 其中 $\theta(x) = (\theta_1(x), \theta_2(x))$ 和 $\theta(y_1, y_2) = y_1/y_2$. 由推论 8.4.14 可以得到 (8.4.40). 若在复合集值映射 $\left\{ \frac{\theta_1}{\theta_2} \right\}(x) = (\psi \circ \theta)(x)$ 中, $\theta(x) =$

(x, x) , $\psi(y_1, y_2) = \theta_1(x)/\theta_2(x)$, 利用推论 8.4.15 即可以得到 (7.4.41). \square

容易推得下面结论.

定理 8.4.23 设集合 $\theta_1(x), \theta_2(x) \subset R (x \in R^n)$, 集值映射 $\theta_1: R^n \rightarrow 2^R$ 和 $\theta_2: R^n \rightarrow 2^R$ 是闭的. 考虑集值映射 $\theta_1 \vee \theta_2: R^n \rightarrow 2^R$, $x \mapsto (\theta_1 \vee \theta_2)(x)$,

$$(\theta_1 \vee \theta_2)(x) = \{z \in R \mid \exists y_1 \in \theta_1(x), y_2 \in \theta_2(x), \\ \max\{y_1, y_2\} = z\}.$$

设 $x \in R^n, z \in (\theta_1 \vee \theta_2)(x)$. 若集值映射 $\Xi: R^n \times R \rightarrow 2^{R^2}$, $(x, z) \mapsto \Xi(x, z)$,

$$\Xi(x, z) = \{(y_1, y_2) \in R^2 \mid y_1 \in \theta_1(x), y_2 \in \theta_2(x), \\ \max\{y_1, y_2\} = z\}$$

在 (x, z) 处是局部有界的, 则对任意的 $\alpha \in R$ 有

$$\partial^M(\theta_1 \vee \theta_2)(x, z)(\alpha) \subset \bigcup_{\substack{(y_1, y_2) \in \Xi(x, z) \\ \alpha \in \Gamma(\alpha)}} [\partial^M(\theta_1, \theta_2) \\ \times (x, y_1, y_2)(\lambda\alpha, (1-\lambda)\alpha)],$$

其中 $\Gamma(\alpha) = [0, 1]$ (当 $\alpha \geq 0$) 和 $\Gamma(\alpha) = \{0, 1\}$ (当 $\alpha < 0$). 再若有

$$\partial^M\theta_1(x, y_1)(0) \cap (-\partial^M\theta_2(x, y_2)(0) = \{0\} \\ \forall (y_1, y_2) \in \Xi(x, z),$$

则

$$\partial^M(\theta_1 \vee \theta_2)(x, z)(\alpha) \\ \subset \bigcup_{\substack{(y_1, y_2) \in \Xi(x, z) \\ \alpha \in \Gamma(\alpha)}} [\partial^M\theta_1(x, y_1)(\lambda\alpha) + \partial^M\theta_2(x, y_2)(1-\lambda)\alpha].$$

参 考 文 献

- [1] Köthe G. Topologische Lineare Räume. Berlin-Götting: Springer-Verlag, 1960
- [2] Jameson G J O. Ordered Linear Space. Lecture Notes in Math, 104, Berlin: Springer-Verlag, 1970
- [3] Rockafellar R T. Convex Analysis. Princeton Mathematics Ser, Vol 28, Princeton: Princeton Univ. Press, 1970
- [4] Clarke F H. Necessary conditions for nonsmooth problems in optimal control and the calculus of variations. Ph. D thesis. Univ. of Washington, 1973
- [5] Ekeland I et Teman R. Analyse Convexe et Problèmes Variationnels. Dunod-Gauthier-Villars, 1974
- [6] Bartle R G. The Elements of Real Analysis. 2nd ed. New York: John Wiley & Sons, 1976
- [7] Пшеничный Б Н. Выпуклый Анализ и Экстремальные Задачи. Москва: Наука, 1980
- [8] Tanino T, Sawaragi Y. Conjugate maps and duality in multiobjective optimization. J. of Opti. Theory and Appl., 1980, 31(4): 473~499
- [9] Ioffe A D. Nonsmooth analysis; differential calculus of nondifferentiable mappings. Trans. Amer. Math. Soc., 1981, 266:1~56
- [10] Канторович Л В. 泛函分析(上册), 刘证等译. 北京: 高等教育出版社, 1982
- [11] Clarke F H. Optimization and Nonsmooth Analysis. New York: John Wiley & Sons-Interscience, 1983
- [12] Köthe G. Topological Vector Spaces I. Berlin: Springer-Verlag, 1983
- [13] Kannippan P. Necessary conditions for optimality of nondifferentiable convex multiobjective programming. J. of Opti. Theory and Appl., 1983, 40(2):167~174
- [14] Minami M. Weak Pareto-optimal necessary conditions in a nondifferentiable multiobjective programming on a Banach space. J. of Opti. Theory and Appl., 1983, 41(3): 451~461

- [15] Jan Van Tiel. *Convex Analysis, An Introductory Text*, John Wiley & Sons, 1984
- [16] Aubin J P. Ekeland I. *Applied Nonlinear Analysis*. New York: John Wiley & Sons, 1984
- [17] Aubin J P. *Lanalyse non Lineaire et ses Motivations Economiques*. Paris: Masson, 1984
- [18] 夏道行, 杨业立. 线性拓扑空间引论. 上海: 上海科学技术出版社, 1986
- [19] 张恭庆, 林源渠. 泛函分析讲义(上册). 北京: 北京大学出版社, 1987
- [20] 夏道行等. 泛函分析第二教程. 北京: 高等教育出版社, 1987
- [21] Keryszing E. 泛函分析导论及应用. 蒋正新译. 北京: 北京航空学院出版社, 1987
- [22] Ferro F. A minimax theorem for vector-valued functions. *J. of Opti. Theory and Appl.*, 1989, 60(1): 19~31
- [23] Luc D T. *Theory of Vector Optimization*. Berlin: Springer Verlag, 1989
- [24] 胡毓达, 胡一凡. 锥拟凸与拓扑向量空间多目标最优化有效解集和弱有效解集的连通性. *应用数学学报*, 1989, 12(1): 115~123
- [25] 史树中. 凸分析. 上海: 上海科学技术出版社, 1990
- [26] 胡毓达, 徐永明. 扰动多目标规划的次微分稳定性. *数学学报*, 1992, 35(5): 577~586
- [27] 胡毓达. 多目标规划有效性理论. 上海: 上海科学技术出版社, 1994
- [28] Morduckhovich B S. Generalized differential calculus for nonsmooth and set-valued mappings. *J. of Math. Anal. and Appl.*, 1994, 183: 250~288
- [29] Lai Jiu Lin. Optimization of set-valued function. *J. of Math. Anal. and Appl.*, 1994, 186(1): 30~51
- [30] Morduckhovich B S. Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis. *Trans. Amer Math Soc.* 1994, 343(2): 609~637
- [31] Tanaka T. Generalized Quasiconvexities, Cone saddle points, and minimax theorem for vector-valued functions. *J. of Opti. Theory and Appl.*, 1994, 81(2): 355~377
- [32] 周性伟. 实变函数. 北京: 科学出版社, 1998

- [33] Hu Yuda, Meng Zhiqing. The existence theorem of the cone-weak subdifferential of set-valued mapping. Appl. Math. A J. of Chinese Univ., 1998, 13B(4):411~415
- [34] 孟志青. 集值函数 Hahn-Banach 定理. 应用数学和力学, 1998, 19(1): 55~61
- [35] 孟志青, 唐勇. 集值函数的一种广义微分的存在性. 应用数学, 1998, 11(4):99~101
- [36] 胡毓达, 孟志青. 拓扑向量空间集值映射的锥有效次微分及其存在性. 上海交通大学学报, 2000, 34(8):1114~1117

- [33] Hu Yuda, Meng Zhiqing. The existence theorem of the cone-weak subdifferential of set-valued mapping. Appl. Math. A J. of Chinese Univ., 1998, 13B(4):411~415
- [34] 孟志青. 集值函数 Hahn-Banach 定理. 应用数学和力学, 1998, 19(1): 55~61
- [35] 孟志青, 唐勇. 集值函数的一种广义微分的存在性. 应用数学, 1998, 11(4):99~101
- [36] 胡毓达, 孟志青. 拓扑向量空间集值映射的锥有效次微分及其存在性. 上海交通大学学报, 2000, 34(8):1114~1117